

Acts, Attitudes, and Rational Control*

Douglas W. Portmore

Abstract: I argue that when determining whether an agent ought to perform an act, we should not hold fixed the fact that she's going to form certain reasons-responsive attitudes (e.g., beliefs, desires, and intentions). For, as I argue, agents have, in the relevant sense, just as much control over which attitudes they involuntarily form as which acts they voluntarily perform. This is important because what effect an act will have on the world depends not only on which acts the agent will simultaneously and subsequently perform, but also on which attitudes she will simultaneously and subsequently form. And this leads me to adopt a new type of practical theory, which I call *rationalist possibilism*. On this theory, we first evaluate the entire set of things over which the agent presently exerts control, where this includes the formation of certain attitudes as well as the performance of certain acts. And, then, we evaluate individual acts as being required if and only if, and because, there is such a set that is itself required and that includes that act as a proper part. This theory has two unusual features. First, it is not exclusively act-orientated, for it requires more from us than just the performance of certain voluntary acts. Additionally, it requires us to involuntarily form certain attitudes. Second, it is attitude-dependent in that it holds that which acts we're required to perform depends on which attitudes we're required to form. I then show how these two features can help us both to address certain puzzling cases of rational choice and to understand why most typical practical theories (utilitarianism, virtue ethics, rational egoism, Rossian deontology, etc.) are problematic.

Keywords: acts, actualism, attitudes, control, deontic logic, obligations, options, possibilism, rationality, reasons-responsiveness, Regan, Ross

To act is to effect a change in the world. And so, when we act, we act with the aim of making the world go a certain way. The aim needn't have anything to do with the causal consequences of the act. The aim could be merely to perform the act in question. I can, for instance, run merely with the aim of running. Nevertheless, if I run, I effect a change in the world: the world becomes one in which I'm running.

But whether an act will effect a certain change in the world often depends on what its agent and others will simultaneously and subsequently be doing. For instance, the commanding officer of a ballistic missile submarine can effect the launch of a missile by turning his key only if the executive officer will simultaneously be turning his key.

* Draft of May 20, 2015. The latest draft can always be found at <http://bit.ly/1Jtb8j>.

Your writing a \$200 check to Oxfam will effect some good only if you're going to mail the check afterwards. And my saying "I never break the law" will have the effect of making me a liar only if I say this with a straight face as opposed to a sly wink and a knowing smile.

Now, agents certainly have reasons for and against effecting such changes in the world and thus reasons for and against performing the acts that would (or could) have these effects. But, as we've just seen, whether these acts would (or could) have these effects depends on what their agents and others will simultaneously and subsequently be doing. So we need to know whether to hold these other acts fixed. And, in general, we need to know what to hold fixed when assessing whether an agent ought to perform a given act.

To illustrate the issue, consider the following case.

Two Drugs: Dr. Singh is deliberating at t_0 about what drugs, if any, to give her patient, Patrick, at t_2 . In fact, there are exactly two drugs available to her: A and B. If she gives Patrick both A and B at t_2 , using her right hand to inject him with A and her left hand to inject him with B, he'll be cured of his painful, but non-fatal, medical condition. If she gives him just one of A and B at t_2 , he'll die immediately. And if she gives him neither A nor B at t_2 , he'll be rendered incurable, having, then, to live the rest of his life with this painful condition. And although this would be quite bad, it is not as bad as his dying immediately. See Table 1. Now, as a matter of fact, Dr. Singh is going to form at t_1 the intention to kill him by giving him just one of A and B at t_2 . For, as it turns out, she is Patrick's jilted lover. Moreover, she's going to follow through with this intention by giving him just A at t_2 . Given that she's going to form, and then follow through, with the intention to give him just one of A and B at t_2 , the following two subjunctive conditionals are true: (SC₁) if she were to give him A at t_2 , she would not give him B at t_2 and (SC₂) if she were to give him B at t_2 , she would not give him A at t_2 . Nevertheless, if she were to respond appropriately to the decisive reason she has at t_0 for curing him and thereby come instead to form at t_1 the intention to cure him by giving him both A and B at t_2 , she would follow through with this intention and Patrick would live happily ever after. Assume that Dr. Singh will lose her medical license if she fails to cure him and will be sent to prison if she kills him.

So Dr. Singh *is* going to give Patrick A at t_2 . But we can still ask whether she *ought* to give him A at t_2 .¹ The answer to this question depends on what we are to hold fixed. If, on the one hand, we are to hold fixed the fact that she's going to form at t_1 the intention to kill him and, thus, would not give him B at t_2 if she were to give him A at t_2 , then the answer seems to be that she should not give him A at t_2 . For given that she's not going to give him B at t_2 , it would be better if she didn't give him A at t_2 —at least, that way he'll live. But if, on other hand, we are not to hold fixed the fact that she's going to respond inappropriately to her reasons and form at t_1 the intention to kill him, then the answer seems to be that she should give him A at t_2 . For if she were to respond appropriately to her reasons, she would form at t_1 the intention to give him both A and B at t_2 and then follow through by doing just that. And she cannot give him both A and B at t_2 without giving him A at t_2 . So she should give him A at t_2 .

Table 1

<i>Two Drugs</i>	Her right hand administers A at t_2 .	Her right hand does not administer A at t_2 .
Her left hand administers B at t_2 .	Patrick is cured.	Patrick dies.
Her left hand does not administer B at t_2 .	Patrick dies.	Patrick lives but is rendered incurable.

In this paper, I'll argue that when assessing whether an agent ought to perform an act x , we should not hold fixed the fact that she's going to y if she has, at present, the same control over whether she y 's as she does over whether she x 's. And I'll argue that

¹ Throughout this paper, I'll be concerned with what agents objectively ought to do (all things considered), where S objectively ought to ϕ if and only if the facts provide S with decisive reason for ϕ -ing. In this sense, a bomb-disposal technician ought to cut the red wire just in case doing so would in fact deactivate the bomb, and this holds even if she justifiably believes that she ought not to cut the red wire and even if all her evidence suggests that cutting the red wire would detonate the bomb. And although I'm interested in what agents *objectively* ought to do as opposed to what they *subjectively* ought to do, we may assume that in my examples the agents know all the relevant facts and, thus, that their objective and subjective obligations coincide. Lastly, although I'm interested in what agents ought to do, *all things considered*, as opposed to what they *morally* ought to do, I have constructed my examples so that their moral obligations coincide with their all-things-considered obligations. Thus, in all my examples, it would be imprudent for the relevant agents to violate their moral obligations.

the relevant sense of ‘control’ is such that agents have the same sort of control over whether they involuntarily form certain reasons-responsive attitudes (such as beliefs, desires, and intentions) as they do over whether they voluntarily perform certain actions.² This, as I’ll show, has two important implications for what sort of practical theory (i.e., a theory about what agents ought to do) we should accept. First, it implies that an adequate practical theory must not be exclusively act-orientated. That is, it must require more from us than just the performance of certain voluntary acts. It must, additionally, require us to involuntarily form certain attitudes. Second, it implies that an adequate practical theory must be attitude-dependent. That is, it must hold that which acts we’re required to perform depends on which attitudes we’re required to form. I then show how these two implications can help us both to address certain puzzling cases of rational choice and to understand why most typical practical theories (utilitarianism, virtue ethics, rational egoism, Rossian deontology, etc.) are problematic.

1. Actualism

On one extreme view, we should hold fixed everything but the fact that an agent is (or is not) going to perform x when assessing whether she ought to perform x . I’ll call this view:

² The notion of a *reasons-responsive attitude* is, perhaps, the same as 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 reason 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.

Binary Actualism: A subject ought to perform an act if and only if she ought to prefer the prospect of her performing that act to the prospect of her not performing that act (Jackson and Pargetter 1986). More precisely, for any time t , subject S , and act-option x for S , S ought, as of t , to x if and only if, and because, S ought, as of t , to prefer $P(x)$ to $P(\neg x)$, where $\neg x$ is the “act” of not performing x , $P(x)$ is the prospect of S ’s x -ing, and $P(\neg x)$ is the prospect of S ’s $\neg x$ -ing.

Before we can assess this view, I need to clarify three things about it. First, we need to talk about the prospect of S ’s x -ing (or $\neg x$ -ing) as opposed to how the world would be if S were to x (or $\neg x$), because there may be no fact of the matter as to how the world would be if S were to x (or $\neg x$). For one, the laws of nature may be indeterministic. And if they are, there may be several different ways that the world *could* turn out rather than just one way that it *would* turn out. For another, x may be underspecified, and S may lack control over how x is to be further specified. And, in that case, there would be no fact of the matter as to how the world would have been had S x -ed. Imagine, for instance, that Cristobal is playing craps and passes on being the shooter. How would the world have been had Cristobal rolled the dice? Would it have been a world in which he rolled boxcars, snake eyes, a hard eight, or what? There may be no determinate answer to this question.³ For what he rolls depends on minute differences in the force and trajectory of his throw, and he lacks the dexterity to determine the precise force and trajectory of his throw. So, for these two reasons, we should, as I have, formulate binary actualism in terms of prospects as opposed to the way the world would be if she were to x (or $\neg x$), where prospects are just probability distributions over the various possible ways the world could be if S were to x (or $\neg x$).

Second, note that, in the above formulation, I talk about what an agent ought, *as of a certain time*, to do. This is because I assume that “ S ought, as of t , to ϕ ” implies that “ ϕ -ing is, as of t , an option for S .”⁴ And, thus, what an agent ought (or is obligated) to do

³ For a defense of this claim, see Hare 2011 and Vessel 2003.

⁴ This assumption has been disputed. Walter Sinnott-Armstrong has, for instance, argued against this assumption—see his 1984. For some of the reasons why I find his arguments unpersuasive, see Streumer 2003. In any case, the reader may note that my conclusions in this paper are conditional upon this assumption.

will change as her options change.⁵ To illustrate, imagine that, last week, Alejandra had the option of enlisting in either the Army or the Navy. But now that she has enlisted in the Army, she can't enlist in the Navy. And supposed that Alejandra had promised her father a month ago that she would enlist in the Navy on her birthday, which is today. A month ago, then, she had an obligation to enlist in the Navy on her birthday. But, as of yesterday, when she enlisted in the Army, she no longer has the option of enlisting in the Navy. Thus, she no longer has an obligation to do so. Of course, she may have an obligation to apologize to her father for breaking her promise. But she cannot now be required to do what isn't even an option for her. This means that, if we are to be careful, we must make explicit the relevant temporal indices when talking about options and obligations.⁶

Third, it's important to note that binary actualism is not committed to consequentialism, for it needn't be that *S* ought to prefer $P(x)$ to $P(\neg x)$ just in case $P(x)$ has more expected value than $P(\neg x)$. It could be, for instance, that *S* ought, as of t , to prefer a world in which she doesn't violate Kant's categorical imperative at t' to any world in which she does violate Kant's categorical imperative at t' ($t \leq t'$). And, in that case, she should prefer the prospect of her performing some act that wouldn't violate the categorical imperative to the prospect of her performing any act that would violate the categorical imperative, and this would be so even if her violating the categorical imperative would have greater expected value. Thus, binary actualism is quite ecumenical.

Having made these clarifications, I will now argue that we should reject binary actualism. It is, as Jacob Ross (2012) has demonstrated, subject to at least two devastating objections. To see this, consider the following example.

⁵ Although I admit that "S ought to ϕ " is not, strictly speaking, equivalent to "S is obligated to ϕ ," I will in this paper use the two phrases interchangeably, because the subtle difference between the two isn't relevant to anything that I'm going to say here.

⁶ As I think of them, temporally-indexed options and obligations refer to properties that are possessed by the agent at specific 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."

The Poisoning: Alice is entertaining a love rival, Betty, who has just asked her for a glass of water. With malice aforethought, Alice is about to give Betty a glass of arsenic solution. When she does, Betty will drink the contents of the glass and die. Of course, Alice has many other options. She could instead give her a glass of water, as she requested. Or she could give her a glass of Drano, or a glass of bleach, or a glass of ammonia. If she were to give her a glass of water, she would drink it and be happy. If she were to give her a glass of Drano, bleach, or ammonia, then while she would not drink enough of the liquid for it to be fatal, she would drink enough to suffer severe and irreversible damage to her esophagus. Assume that Alice will receive a longer prison sentence if she gives Betty a glass of arsenic solution than if she gives her a glass of Drano, ammonia, or bleach and that she'll stay out of prison altogether if she gives her a glass of water.⁷

Binary actualism implies that Alice ought to give Betty a glass Drano, for, given the stipulations of the case, the prospect of Alice's giving Betty a glass of Drano is a 100% chance of Betty's suffering severe and irreversible damage to her esophagus, and this is preferable to the prospect of Alice's not giving Betty a glass of Drano, which is a 100% chance of Betty's dying of arsenic poisoning. This is the prospect of Alice's not giving Betty a glass of Drano, because the world that's nearest to the actual world that involves Alice's not giving Betty a glass of Drano is the actual world itself, a world in which she gives Betty a glass of arsenic solution.⁸ But the idea that Alice ought to give Betty a glass of Drano is absurd. She should not give Betty a glass of Drano, but should instead give her a glass of water, which is just as much an option for her as giving her a glass of Drano.

What's more, binary actualism implausibly implies that Alice ought to do several incompatible things. Not only should she give Betty a glass of Drano, but she should give her a glass of bleach, and she should give her a glass of ammonia. In each case, the prospect of Alice's giving Betty a glass of that liquid is preferable to the prospect of her not doing so, which is a 100% chance of Betty's dying of arsenic poisoning. But it's ab-

⁷ This is a revised version of Ross's case entitled *Arsenic and Old Ace*—see his 2012, p. 75.

⁸ The world in which she gives Betty a glass of water is not as close to the actual world as the actual world is to itself, for the world in which she gives Betty a glass of water would be one in which she either does not want to kill Betty or is not as diligent as she is in the actual world about making sure she fills the glass with the intended liquid. For more on how to evaluate the relevant sorts of subjunctive conditionals, see Vessel 2003.

surd to hold that Alice should do all these things, for it could be, as we'll indeed suppose, that she can't do all these things. For assume that she has only one glass. The problem, then, is that binary actualism is committed to an explosion of incompatible obligations. If there are a million options that are less bad than what Alice actually does, then binary actualism implies that she's obligated to perform all of them even if she can't perform all of them. Thus, binary actualism not only implies that Alice should do horrible things but also that Alice should do incompatible things. We should, therefore, reject binary actualism.

There is, however, a close cousin to binary actualism that avoids these two problems:

Multitudinous Actualism: A subject ought to perform an act if and only if she ought to prefer its prospect to that of any available alternative. More precisely, for any time t , subject S , and act-option x for S , S ought, as of t , to x if and only if, and because, for every alternative act-option y , S ought, as of t , to prefer $P(x)$ to $P(y)$.

Unlike binary actualism, this view does not require us to hold fixed the fact that Alice would give Betty a glass of arsenic solution if she were to refrain from giving her a glass of Drano. Instead, we are to consider the possibility that Alice gives Betty a glass of water. Accordingly, multitudinous actualism does not imply that Alice ought to give Betty a glass of Drano. Nor does it imply that Alice should do many incompatible things. It doesn't, for instance, imply that Alice should give Betty a glass of Drano and a glass of bleach and a glass of ammonia. It implies only that she should give Betty a glass of water. Thus, multitudinous actualism is not as implausible as binary actualism.⁹ But it is still implausible, for it entails that both of the following two plausible principles are false:

⁹ Although binary actualism has some problems that multitudinous actualism doesn't have, binary actualism has all the problems that multitudinous actualism has. I won't be explicit about this in the body of the paper, but all the problems that I'm about to attribute to multitudinous actualism are shared by binary actualism.

Inheritance (INH): If a subject has both the option of x -ing and the option of y -ing but cannot x without y -ing, then “S ought to x ” entails “S ought to y .” More precisely, for any subject S, acts x and y , and times t , t' , and t'' ($t \leq t' \leq t''$), if x -ing at t' and y -ing at t'' are, as of t , both options for S but x -ing at t' without y -ing at t'' is not, then S’s being, as of t , obligated to x at t' entails that S is, as of t , obligated to y at t'' .¹⁰

Circumstances (CIR): In the determination of whether a subject ought to x (or $\neg x$), we are not to hold fixed, as part of her circumstances, the fact that she’s going to y (or $\neg y$) if she has, at present, the same control over whether she y ’s as she does over whether she x ’s. More precisely, for any subject S, acts x and y , and times t , t' , and t'' ($t \leq t' \leq t''$), in the determination of whether S ought, as of t , to x (or $\neg x$) at t' , we are not to hold fixed, as part of S’s circumstances, the fact that S is going to y (or $\neg y$) at t'' if S’s $\neg y$ -ing (or y -ing) at t'' is, as of t , just as much an option for S as S’s x -ing (or $\neg x$ -ing) at t' is.

I’ll argue that we should accept both INH and CIR and, so, should reject multitudinous actualism, which entails that both are false. Let me start with why we should accept INH. One reason to accept INH is that it allows us to explain what has been called *coarseness*: “S ought to ϕ ” can be true even though there are many instances of S’s ϕ -ing that are impermissible.¹¹ Consider, for instance, that although it would be impermissible for me to feed my daughter poisoned food, too much food, or shards of glass, the propo-

¹⁰ INH is similar to standard deontic logic’s rule of monotonicity (RM), which holds that, for any two propositions p and q , if p entails q , then ‘it ought to be that p ’ entails ‘it ought to be that q ’. But unlike RM, INH is not subject to either the Good Samaritan Paradox or the Paradox of Epistemic Obligation. It is, however, subject to Ross’s Paradox. Suppose that Kimo ought to mail a letter. Because Kimo cannot mail the letter without either mailing or burning it, it follows from INH that Kimo ought to either mail or burn the letter. But many find it odd to say that Kimo is under an obligation to either mail or burn the letter. The standard response, of course, is to claim that this only sounds odd because it violates Grice’s maxim of quantity, which implies that saying that Kimo ought to either mail or burn the letter when we know that he ought to mail the letter is inappropriately uninformative—see, for instance, Wedgwood 2006. In any case, the claim that Kimo is under an obligation to either mail or burn the letter will have implausible practical implications only if we hold both that, by burning the letter, Kimo can fulfill his obligation to either mail or burn the letter and that the mere fact that his burning the letter would fulfill this obligation is a reason for him to do so. But I think that we have independent reasons for rejecting this conjunction of claims. And if we reject this conjunction, we can accept that Kimo is under an obligation to either mail or burn the letter while denying that he has any reason to burn the letter. For more on these paradoxes, see McNamara 2014.

¹¹ See Cariani 2013. Unlike Cariani, I don’t take at face value some of the intuitions that Cariani thinks speak against INH, such as the one concerning Ross’s Paradox or the one concerning Professor Procrastinate. See above and below, respectively.

sition “I ought to feed my daughter” is true. It’s true because it is entailed by the proposition “I ought to feed my daughter a proper diet” and I cannot feed her a proper diet without feeding her. Thus, as INH would have it, the proposition “S ought to y ” is true so long as there is some x such that S ought to x but cannot x without y -ing.

Another reason to accept INH is that it allows us to give a systematic explanation as to why all the following seemingly valid inferences are indeed valid:

- (I1) Ibrahim ought to bake a cake. But he can’t bake a cake without turning on the oven (because his turning on the oven is a necessary means to his baking a cake). Therefore, he ought to turn on the oven.
- (I2) Irene ought to marry a bachelor. But she can’t marry a bachelor without marrying an unmarried man (because a bachelor just is an unmarried man). Therefore, she ought to marry an unmarried man.
- (I3) Isabella ought to drive under the speed limit, which is 55 mph. But she can’t drive under 55 mph without driving under 75 mph (because her driving under 55 mph is an instance of her driving under 75 mph). Therefore, she ought to drive under 75 mph.
- (I4) Isaac ought to write the check and then mail it. But he can’t write the check and then mail it without writing the check (because his writing the check is a proper part of his writing the check and then mailing it). Therefore, he ought to write the check.¹²

If we accept multitudinous actualism, we must eschew this very plausible explanation for the validity of I1–I4. This is because multitudinous actualism implies that INH is false. To see this, consider the following case.

Professor Malice: He receives an email at t_0 from a journal inviting him to review a book. He is the best person for the job, and he has the time to do it. The best thing, then, would be for him to accept the invitation at t_1 and then write a fair review when the book arrives at t_{100} . However, if Professor Malice were to accept,

¹² We could subsume I2–I4, but not I1, under the following pattern (I2*): “S ought to perform x . But S can’t perform x without performing y (because her performing x logically necessitates her performing y). Therefore, S ought to perform y .”

he would not write a fair review. This is not because he's weak-willed or anything like that. Indeed, he'll do whatever he intends to do. But given his animosity towards the book's author, he's going to form at t_1 the intention to write an unfair review. Thus, although the best thing would be for him to accept and then write a fair review, what would in fact happen if he were to accept is that he would write an unfair review. And his accepting and writing an unfair review is the worst thing that could happen, as this would undermine the book's chances of receiving the attention that it deserves as well as hurt Professor Malice's own reputation. It would be better, then, for him to decline the invitation and let the next best person write a fair review. The total payoffs are as depicted in Table 2, and Professor Malice's share is, in each case, proportionate to the totals.¹³

Table 2

<i>Professor Malice...</i>	<i>...accepts at t_1.</i>	<i>...does not accept at t_1.</i>
<i>...writes a fair review at t_{100}.</i>	100 utiles	-10 utiles ¹⁴
<i>...does not write a fair review at t_{100}.</i>	-100 utiles	75 utiles

Multitudinous actualism denies that Professor Malice ought to accept the invitation (and let's call the act of accepting ' a '). For it is not the case that he ought to prefer the prospect of his accepting to that of every available alternative act. He should not, for instance, prefer the prospect of his accepting (i.e., $P(a)$) to the prospect of his not-accepting (i.e., $P(\neg a)$). After all, he ought, in this type of situation, to prefer the prospect with the greatest expected value. And, given the stipulations of the case, $P(a)$ is a 100% chance of the worst outcome (the one in which Professor Malice's publishes an unfair

¹³ This is a variation on Jackson and Pargetter's case of Professor Procrastinate—see their 1986. But my case is importantly different from theirs. In my case, it's stipulated that if Professor Malice were, upon receiving the invitation, to form the intention to accept and write a fair review, he would then follow through on this intention and later write a fair review. By contrast, it seems that Jackson and Pargetter are imagining that Professor Procrastinate would not follow through on his intention to write the review even if he were, upon receiving the invitation, to form the intention to accept and write.

¹⁴ This is negative, because, in this scenario, he wastes his time writing a review that will never be published.

review), whereas $P(\neg a)$ is a 100% chance of the second best outcome (the one in which the journal gets the next best person to write a fair review).

But even though multitudinous actualism denies that Professor Malice ought to accept, it holds that he ought to accept and write a fair review (and let's call the act of writing a fair review ' f '). For Professor Malice ought to prefer $P(a \& f)$ to the prospect of his performing any alternative to $(a \& f)$.¹⁵ After all, the only alternatives are: (1) his not accepting ($\neg a$); (2) his not accepting and writing a fair review that will never be published ($\neg a \& \neg f$); (3) his not accepting and not writing a fair review either by not writing any review or by writing an unfair review that will never be published ($\neg a \& \neg f$); and (4) his accepting and not writing a fair review either by not writing any review or by writing an unfair review that will be published ($a \& \neg f$). And Professor Malice ought to prefer $P(a \& f)$ to the prospects of any of these alternatives. So multitudinous actualism commits us to the view that he ought to accept and write but that it is not the case that he ought to accept. Yet he cannot accept and write without accepting. Thus, multitudinous actualism commits us to the denial of INH. And since we should accept INH so that we can explain coarseness and the validity of I_1 – I_4 , we should reject multitudinous actualism.

Admittedly, the fact that multitudinous actualism commits us to the denial of INH will not dissuade proponents of actualism. These proponents are well aware that their view commits them to the denial of INH, and they are willing to countenance this.¹⁶ But it seems that even they must admit that there is a cost to rejecting INH. For we often make inferences such as I_1 – I_4 , which all seem to be valid. Moreover, coarseness seems both true and in need of explanation. So, other things being equal, it would be better to accept INH so that we can explain both the truth of coarseness and the validity of I_1 – I_4 . Proponents of multitudinous actualism just don't think that other things are equal. They believe that multitudinous actualism does a better job of accounting for our intuitions

¹⁵ S 's ψ -ing is an alternative to S 's ϕ -ing if and only if S has the option of ψ -ing, the option of ϕ -ing, but not the option of both ψ -ing and ϕ -ing.

¹⁶ See, for instance, Jackson 1988. And note that I'm claiming, contrary to Jackson, that the t_0 -temporal-part of Professor Malice, and not just that the temporally-extended Professor Malice, ought, as of t_0 , to accept the invitation at t_1 .

than any alternative view does. In what follows, I'll call that belief into question. In any case, though, this is not my only argument against multitudinous actualism.

Another reason to reject multitudinous actualism is that it implies \neg CIR. CIR is, I believe, quite plausible on its face. But to pump your intuitions, consider that, in the determination of whether a subject ought to x (or $\neg x$), we do not hold fixed, as part of her circumstances, the fact that she's going to x (or $\neg x$). Nor would we hold fixed any further facts stemming from this one. The reason being that such facts are true only in virtue of the way the subject is going to exert her control. To illustrate, suppose that a criminal named Katia is going to kill a police officer named Lazar in order to take his gun. Now, it would be absurd to argue that Katia ought to kill Lazar and take his gun, because, given that she's going to kill Lazar, Lazar's brother, Leo, is going to come after her, and so she'll need Lazar's gun to protect herself from him. It's absurd to hold fixed the fact that Leo is going to come after her when assessing whether she should kill Lazar, because the only reason it's a fact that Leo is going to come after to her is because she's going to kill Lazar. And when assessing how an agent should exert her control, we should not hold fixed the way she is in fact going to exert her control or any of the other facts that are just as much under her control as this one. So we should not hold fixed the fact that a subject is going to y (or $\neg y$) when assessing whether she ought to x if she has the same control over whether she y 's as she does over whether she x 's?

To further pump your intuitions, imagine that it's t_0 and Dr. Singh is deliberating about what she ought to do at t_2 . It seems that she should not reason as follows.

Bad Reasoning: (B1) I ought to refrain from giving Patrick A at t_2 if I'm not going to give him B at t_2 . (B2) As a matter of fact, I'm not going to give him B at t_2 . Therefore, (B3) I ought to refrain from giving him A at t_2 .

Yet this is exactly how the actualist argues that she should reason (see Jackson and Pargetter 1986, 238–239).¹⁷ But, to see that such reasoning is fallacious, we need only consider how Katia might employ the same reasoning.

More Bad Reasoning: (MB₁) I ought to kill Lazar and take his gun if Leo is going to come after me. (MB₂) As a matter of fact, Leo is going to come after me. Therefore, (MB₃) I ought to kill Lazar and take his gun.

But, obviously, this reasoning is fallacious. Even granting both MB₁ and MB₂, MB₃ does not follow.¹⁸ Clearly, Katia should not kill Lazar and take his gun. The problem with such reasoning, it seems to me, is that in reasoning on the basis of a premise such as B₂ or MB₂, the subject treats as fixed and unalterable a fact that she is instead able to alter. Take, for instance, B₂. B₂ would be false if Dr. Singh were going to give Patrick B at t_2 . And, as of t_0 , Dr. Singh has the same control over whether she gives Patrick B at t_2 as she does over whether she gives him A at t_2 . In both cases, whether she gives Patrick that drug at t_2 just depends on whether she forms the intention at t_1 to do so. And when engaging in practical reasoning for the sake of forming some intention about what to do, one should not presuppose that one is going to form one intention and not another. So, it seems that Dr. Singh should instead reason as follows.

Good Reasoning: (G₁) I ought to cure Patrick. (G₂) I cannot cure him without giving him A at t_2 . Therefore, (G₃) I ought to give him A at t_2 .

¹⁷ I agree with Jackson and Pargetter (1986, 239) that there is nothing fallacious in reasoning as follows: (1) I ought to arrange for someone to give me a ride home tonight if I'm going to freely choose to drink too much tonight. (2) As a matter of fact, I'm going to freely choose to drink too much tonight, and this is not under my present control, as I will freely choose to drink too much tonight no matter what I intend to do now and, indeed, no matter how I presently respond to my reasons. Therefore, (3) I ought to arrange for someone to give me a ride home tonight. But this does not show that what I call *Bad Reasoning* is unproblematic. For whether Dr. Singh is going to give Patrick B at t_2 depends entirely on how she presently responds to her reasons. So Jackson is correct to claim that that what I will, in fact, do in the future can be relevant to what I ought to do presently even if the future act will be one that I freely choose to perform. But I deny that it is relevant when what I'm going to do in the future depends entirely on how I presently respond to my reasons.

¹⁸ It may seem that denying the validity of this inference commits us to denying the validity of *modus ponens*. But it does so only if we assume that MB₁ is a material conditional with "Leo is going to come after me" as its antecedent and "I ought to kill Lazar and take his gun" as its consequent. And, of course, I deny this assumption, as do many others—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 ought has wide-scope over the entire conditional (e.g., Broome 2004).

And if she were to reason in this way, recognizing and responding appropriately to the superior reasons that she has for curing him, she would then form at t_1 the intention to cure him by giving him both A and B at t_2 . And, in that case, B₂ would be false. So it seems, as CIR would have it, that in determining whether to give Patrick A at t_2 Dr. Singh should not hold as fixed and unalterable the fact that she is not going to give him B at t_2 , when this is just as much under her present control as whether she gives him A at t_2 .¹⁹ Thus, I think that we should accept CIR and reject multitudinous actualism.

Lastly, we should reject multitudinous actualism because it has counterintuitive implications in *Two Drugs*. (It seems to me that it also has counterintuitive implications in *Professor Malice*.) In *Two Drugs*, multitudinous actualism implies that Dr. Singh ought, as of t_0 , to give Patrick both A and B at t_2 but that it is not the case that Dr. Singh ought, as of t_0 , to give Patrick A at t_2 . Yet giving Patrick A at t_2 is an essential part of his giving Patrick both A and B at t_2 , which is the best of all the actions available to Dr. Singh as of t_0 . Even Frank Jackson is no longer willing to countenance such a set of verdicts. He now thinks that in such a case Dr. Singh ought, as of t_0 , to give Patrick A at t_2 on the grounds that this is an essential part of what is best, out of the actions available to her as of t_0 .²⁰

2. Possibilism

I've argued that we should reject actualism. But what should we accept instead? I believe that it's *possibilism*.

Possibilism, unlike actualism, distinguishes maximal options from non-maximal options. An option ' ϕ ' is a maximal option for S if and only if there is no other option ' ψ ' for S such that S ψ -ing logically necessitates S's ϕ -ing, and an option ' ϕ ' is a non-

¹⁹ If the fact expressed by B₂ were instead unalterable in the sense that Dr. Singh was going to give Patrick B at t_2 no matter how she were now to respond to her reasons (that is, not matter what she were now to do, intend, feel, or believe), then we could infer B₃ from B₁ and this unalterable fact. This kind of inference pattern is known as *restricted factual detachment*, whereas the inference to B₃ from B₁ and B₂ is known as (unrestricted) *factual detachment*.

²⁰ See Jackson 2014, pp. 645–646. This represents a significant change from his earlier position—see Jackson and Pargetter 1986, 241.

maximal option for S if and only if it is not a maximal option.²¹ So walking at t_1 won't be a maximal option if walking fast at t_1 is an option. For walking fast at t_1 logically necessitates walking at t_1 . 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, for walking fast at t_1 and then jogging at t_2 logically necessitates walking fast at t_1 .

Possibilism distinguishes between these two types of options, because it evaluates them differently. It evaluates a maximal option in terms of whether the agent ought to prefer its prospect to that of any alternative maximal option. But instead of evaluating non-maximal options in the same way (as actualism does), it evaluates a non-maximal option in terms of whether its performance is necessitated by the performance of some maximal option that the agent ought to perform. Thus, we get

Possibilism: Whereas a subject ought to perform a maximal option if and only if she ought to prefer its prospect to that of any alternative maximal option, a subject ought to perform a non-maximal option if and only if its performance is logically necessitated by the performance of some maximal option that she ought to perform. More precisely, for any time t , subject S, and maximal option μ_1 for S, S ought, as of t , to μ_1 if and only if, and because, for every alternative maximal option μ_2 , S ought, as of t , to prefer $P(\mu_1)$ to $P(\mu_2)$. And for any time t , subject S, and non-maximal option ν for S, S ought, as of t , to ν if and only if, and because, there is a maximal option μ for S such that S ought, as of t , to μ and S's μ -ing logically necessitates S's ν -ing.

Unlike actualism, possibilism is compatible with INH. Indeed, like INH, possibilism entails I2–I4. Admittedly, possibilism doesn't entail I1. For the sort of necessity that is involved in an act's being a necessary means to another is weaker than the sort of necessity involved in the performance of one act logically necessitating the performance of another. But even though possibilism doesn't entail I1, it is compatible with it and, thus, also with INH.

²¹ With respect to the option of ϕ -ing, ψ counts as an *other* option if and only if ϕ and ψ are not identical options. Thus, the option to both x and y is an option other than the option to x . It's just not an *alternative* option.

And possibilism is compatible with CIR. Consider, for instance, *Two Drugs*. Possibilism holds that, in determining whether Dr. Singh ought to give Patrick A at t_2 , we do not hold fixed, as part of her circumstances, the fact that she is not going to give him B at t_2 , because, on possibilism, determining whether she ought to give him A at t_2 involves looking at all of her maximal options, several of which involve her giving him B at t_2 . Thus, we are to hold fluid, not fixed, Dr. Singh's refraining from giving Patrick B at t_2 . Indeed, on possibilism, we conclude that she ought to give him A at t_2 , for all the best maximal options open to her involve her giving him both A and B at t_2 , and her performing such a maximal option logically necessitates her giving him A at t_2 .

Possibilism is more of a family of views than a specific view, so I need to say something about which version of possibilism we should accept. We get different versions of possibilism depending both on what we take an agent's reasons for preferring one prospect to another to be and on what we take an agent's options to be. In this paper, I'll focus on the latter issue.

Traditionally, possibilists have endorsed the following account of options.

Schedulism: A subject has the option of ϕ -ing if and only if she would ϕ if she were to have the right intentions at the right times. More precisely, for any event ϕ , subject S, and times t , t' , t'' , and t''' ($t < t' < t'' < t'''$), S has, as of 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'' .²²

When we combine schedulism with possibilism, we get what I call *schedulist possibilism*. We should, I think, reject schedulist possibilism. To see why, consider the following example.

²² This is adapted from Ross 2012. Schedulists include Feldman 1986 and Zimmerman 1996 (46).

The Safe: A thousand lives will be saved if and only if Safar opens a safe by entering the correct seven-digit combination—viz., 8675309—on his first try. Assume that Safar is one of the thousand whose lives would be saved. And assume (1) that if Safar were to intend at t_0 to enter 8 at t_1 , he would succeed in doing so, (2) that if, after that, he were to intend at t_2 to enter 6 at t_3 , he would succeed in doing that, and (3) that similar things hold for all the remaining digits in the combination. In other words, assume that Safar's entering the correct seven-digit combination on his first try is, as of t_0 , a *schedulist option* for him. Yet I'll stipulate that no matter what Safar's actions, beliefs, desires, and intentions are at t_0 , he's not going to enter the correct combination. For one, he isn't so lucky as to just stumble upon the correct combination. And, for another, he wouldn't enter the correct combination even if he were told at t_0 what it was. For he is, we'll assume, incapable of holding more than one single-digit number in his mind at any given time, and no one is going to tell him at each moment which single-digit number to enter.²³

In this case, schedulist possibilism implies that Safar has, as of t_0 , the option of opening the safe and that, given this and the superiority of its consequences, he is, as of t_0 , obligated to open the safe. I find both implications absurd. Note that we're talking about what is, *as of* t_0 , both an option and an obligation for him, and it's stipulated that no matter what he does, thinks, feels, or intends as of t_0 he won't enter the correct combination. So I don't see how his opening the safe could be either an option or an obligation for him *as of* t_0 . And it's not just that he doesn't know what the correct combination is. For it's stipulated that even if he were told at t_0 what the combination is, he's incapable of translating this momentary knowledge into the required extended course of action, because he's incapable of holding more than one single-digit number in his mind at any given time.

Thus, *The Safe* is importantly different from a case in which, say, there are ten numbered buttons before Safar and he isn't going to push the one button that would save the thousand (assume that it's the one labeled "7") because he's only allowed to

²³ Likewise, schedulist possibilism holds that we are all required to do a great many things that it seems that we cannot do in the sense that's relevant to what we ought to do. Examples include typing out the cure for cancer, writing the next *King Lear*, and defeating grandmaster Anatoly Karpov at chess—see Wiland 2005 and Howard-Snyder 1997.

push one button and he has no idea which is the right one. (Assume that he'll guess that it's the button labeled with his lucky number, which is 3.) In this case, he's not going to save the thousand because he doesn't know which button to push. But, in *The Safe*, he's not going to save the thousand because, even if he had the relevant knowledge at t_0 , he's incapable of translating that into the required, temporally-extended course of action. So my objection to schedulist possibilism isn't simply a version of the kind of epistemic objection that is typically leveled against objective consequentialism. I'm not arguing that it is a mistake to hold that an agent is obligated to perform her best option when she doesn't know which of her options is best. Rather, I'm arguing that it is a mistake to hold that an agent has, as of t_0 , the option of performing a certain sequence of actions when she won't perform that sequence no matter what she does, thinks, feels, or intends at t_1 .

In any case, this isn't my only objection to schedulist possibilism. Consider the following case.

Professor Procrastinate: He will undergo at t_2 some psychological-continuity-preserving but potentially identity-destroying process—pick some process that you take to have these two features whether it be fission, teletransportation, or something else. And let's call the person who enters into this process 'Pre' and the person who exits from it 'Post' (or if there's more than one person who exits from it, let's call the one closest to the North Pole 'Post'). Thus, I want to leave open both the possibility that Pre and Post are numerically identical such that 'Pre' and 'Post' are just two different names that refer to the same person (viz., Professor Procrastinate) and the possibility that Pre and Post are numerically distinct such that 'Pre' and 'Post' refer to distinct persons, both of whom have the same tendency to procrastinate. We are to assume, then, that the process is one where although it's clear that Post is psychologically continuous with Pre, it's unclear whether Post is the same person as Pre. With this in mind, I can fill in the rest of the details. Pre receives at t_0 an email from a journal inviting him to review a book. Given his knowledge and expertise, there is no one *else* (besides Post, if he's numerically distinct) who would do nearly as good a job reviewing the book as he would. The best thing, then, would be for Pre to accept the invitation at t_1 and for Post to write the review, starting when the book arrives at t_{100} . However, even if Pre were to accept at t_1 , Post would never write the review. He is, after all, a procrastinator. But assume that although Post is a procrastinator

who constantly puts off starting projects, he does finish whatever he starts. Thus, he would finish the review if he were to start working on it when the book arrives at t_{100} . But, unfortunately, no matter what Pre does, thinks, feels, or intends at t_1 , Post is never going to start work on the review. When the book arrives at t_{100} , he's going to decide to put it off a week. And when that week arrives, he's going to decide to put it off for another week, and so on and so forth, procrastinating indefinitely. And Post's putting it off indefinitely, never writing the review, is the worse possible thing to have happen—both for the author and for Pre/Post. It would be better, then, if Pre were to decline the invitation at t_1 , in which case the journal would get the next best person to write the review.

According to schedulist possibilism, whether Pre should accept the invitation at t_1 depends on whether Post is numerically identical to Pre. For, given schedulism, Pre's accepting the invitation at t_1 and then writing the review when the book arrives at t_{100} counts as an option for her as of t_0 if and only if Pre and Post are numerically identical. (After all, the variable 'S' in the formulation of schedulism refers to the same person throughout.) So if, on the one hand, accepting and writing counts as an option for Pre as of t_0 , then she is obligated, as of t_0 , to accept the invitation at t_1 . But if, on the other hand, accepting and writing doesn't count as an option for Pre as of t_0 (because Pre and Post are not identical), then she is obligated, as of t_0 , to decline the invitation at t_1 . This is curious. For what seems to matter is not whether Pre and Post are numerically identical, but whether Pre can effect Post's writing the review. After all, what we care about is whether the book is reviewed and whether the best person (*viz.*, Post) writes the review. And so what matters is whether Post will write the review. And, unfortunately, no matter what Pre does, thinks, feels, or intends, Post won't write the review. So another reason to reject schedulist possibilism is that it seems to be more concerned with whether Pre is numerically identical to Post than with whether Pre can effect Post's writing the review by, say, forming the intention to start working on the review when the book arrives at t_{100} .²⁴

²⁴ Christopher Woodard (2009) was the first to point out how schedulist possibilism insists that agents treat future others and their future selves asymmetrically.

Both of the problems that I've raised with schedulist possibilism point to schedulism's conception of an option being too broad. We might think, then, that a much narrower conception is called for, one where S has an option now of doing something later only if S could now form an effective intention to do that something later. More precisely, the idea is this:

Voluntarism: A subject has the option of ϕ -ing if and only if she has the capacity to intend to ϕ and whether she ϕ s just depends on whether she so intends. More precisely, for any ϕ , subject S, and times t , t' , and t'' ($t < t' < t''$, and although t' must immediately follow t , t'' can be any time subsequent to t'), S has, as of t , the option of ϕ -ing at t'' if and only if, and because, S has, as of t , volitional control over whether or not she ϕ s at t'' —that is, she has the capacity to intend (or will) at t' to ϕ at t'' and whether she ϕ s at t'' just depends on whether she so intends (or wills) at t' .

In *The Safe*, there is no intention that Safar could have now that would lead to his subsequently entering the correct combination. And, in *Professor Procrastinate*, there is no intention that that Pre could have now that would result in Post's writing the review. So, on voluntarism, these acts do not count as options as of the present despite the fact that these acts would be performed if their respective agents were to form the right intentions at the right times. Thus, voluntarist possibilism (which is the view we get from combining voluntarism with possibilism) avoids the counterintuitive implications associated with schedulist possibilism.

Nonetheless, we should reject voluntarist possibilism on the grounds that voluntarism is too narrow. It holds that we don't have the option of involuntarily forming various reasons-responsive attitudes, such as beliefs, desires, and intentions. This seems mistaken. Even if an unconscious patient doesn't have the option of, say, forming the intention to raise her arms (because she lacks the capacity to form that intention), I do. After all, I just did a moment ago. I was feeling uncomfortably stiff, realized that I could relieve some of this stiffness by raising my arms in a stretch, and, as a result, formed the intention to raise my arms in a stretch. I formed this intention involuntarily in response

to the reason that I had for doing so. I certainly didn't form this intention voluntarily by intending to form this intention.²⁵ And just as I formed this intention involuntarily in response to my reasons, I often form other reasons-responsive attitudes involuntarily in response to my reasons. For instance, earlier today I formed the belief that I have plenty of milk. For I opened the fridge and saw a full gallon of milk inside. I admit, of course, that I voluntarily opened the fridge and that this resulted in my forming the belief that I have plenty of milk. But it was the act of opening the fridge, not the formation of the belief, that was under my volitional control.

Denying that forming such an attitude counts as an option comes at a steep price. Since "I ought to ϕ " implies " ϕ -ing is an option for me," we would, then, have to deny that I ever ought to form such an attitude. But if I'm aware that there's a full gallon of milk in the fridge, I ought to believe that I have plenty of milk. And if I want to relieve some of my stiffness and believe that I can do so by raising my arms in a stretch, then I ought to intend to raise my arms. So we need an account of options that, unlike voluntarism, allows that just before I had raised my arms in a stretch a few minutes ago I had the option of involuntarily forming the intention to do so. The following provides just such an account.

Rationalism: A subject has the option of ϕ -ing if and only if she has rational control over whether or not she ϕ s. More precisely, for any ϕ , subject S, and times t and t' ($t < t'$), S has, as of t , the option of ϕ -ing at t' if and only if, and because, S has, as of t , rational control over whether or not she ϕ s at t' .²⁶

²⁵ Consider also 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 neither case will I have any reason to drink toxin. Indeed, 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.

²⁶ Note that the sense of 'option' with which I'm concerned is the one for which "S ought (or is obligated), as of t , to ϕ " necessarily implies " ϕ -ing is, as of t , an option for S." Thus, it is no objection to rationalism that it entails that although I just hiccupped, my hiccupping was never an option for me. To claim that my hiccupping was ever an option for me despite my never having had any control over whether I would hiccup is, I believe, to commit a category mistake. Because I

Of course, this raises the question: What is it for S to have, as of t , rational control over whether or not she ϕ s at t'' ? And although I don't have a fully worked-out answer, I owe the reader at least a tentative proposal. I'll start, though, by explaining the general thoughts that have guided me in coming up with this proposal. First, for S to have, as of t , rational control over whether or not she ϕ s at t'' , she must have the sort of control that normal adult human beings typically have over their beliefs and intentions—the sort of control in virtue of which we hold people accountable, not only for their voluntary actions, but also for their beliefs and intentions. So rational control is not equivalent to volitional control, for although I have volitional control over my actions, I don't (at least, not typically) have volitional control over my beliefs and intentions. Second, if S would not ϕ at t'' no matter what acts and attitudes she were to form and perform at t' , then S does not have, as of t , rational control over whether she ϕ s at t'' and so does not have, as of t , the option of ϕ -ing at t'' . Third, for S to have, as of t , rational control over whether or not she ϕ s at t'' , the mechanism by which things of ϕ 's type are produced must be generally responsive to reasons.²⁷ Thus, for S to have rational control over her beliefs, her belief-forming mechanism must be generally responsive to the relevant sorts of reasons.

With these guiding thoughts in mind, I offer the following tentative proposal.

Rational Control: S has, as of t , rational control over whether or not she ϕ s at t'' if and only if she has, as of t , the capacity to respond appropriately at t' to the relevant sorts of reasons and whether or not she ϕ s at t'' just depends on whether and how she responds at t' to her reasons ($t \leq t' < t''$).²⁸

lack control over whether I hiccup, it is not the sort of thing that can be obligatory and, thus, is not the sort of thing that can, in the sense that concerns us here, count as an option.

²⁷ For what it might be for a mechanism to be at least generally/moderately reasons-responsive, see Fischer & Ravizza 1998, pp. 62–91 and especially 243–244. I'm not, however, wedded to their particular account.

²⁸ By the relevant sorts of reasons, I have in mind some class of reasons, such as *moral* reasons or *prudential* reasons. A subject may have the capacity to respond to one class of reasons but not to another. For instance, it may be that a certain psychopath is incapable of recognizing that the fact that killing gently causes less suffering is a reason to kill gently. And because of this, it may be that his awareness of the fact that killing gently causes less suffering could not lead him to kill gently. And, thus, he would lack the capacity to respond appropriately to such moral reasons. This sort of case is very different from one in which some malicious person fails to respond appropriately to the moral reasons that he has for killing gently despite having the capacity to do so.

Furthermore, I propose that S has, as of t , the capacity to respond appropriately to the relevant sorts of reasons if and only if, and because, S has some generally reasons-responsive mechanism by which things of ϕ 's type are produced and this mechanism is, as of t , (1) operational, (2) attuned to the relevant reasons-constituting facts, and (3) unfettered in its normal operation. Condition 1 rules out the possibility that a subject whose relevant attitude-forming mechanism is inoperative due to unconsciousness or similar mental incapacitation could have the option of forming the relevant attitudes. Condition 2 rules out the possibility that a subject whose relevant attitude-forming mechanism is not attuned to the relevant reasons-constituting facts could have the option of forming the relevant attitudes in response to facts to which she lacks epistemic access. And condition 3 rules out the possibility that a subject whose relevant attitude-forming mechanism is being manipulated so that it is incapable of forming the attitudes that it would normally form has the option of forming these attitudes. But keep in mind that this is all very tentative. The only thing that I'm absolutely committed to here is that there is a type of control that we have over our beliefs, intentions, and other reasons-responsive attitudes and in virtue of which it is appropriate to hold us responsible for them despite our lacking volitional control over them. Rational control is just whatever this is, regardless of whether I have the details correct.

Now, if, as I suggest, we should accept rationalism as opposed to voluntarism, we must think that it is rational control as opposed to volitional control that matters in determining what our obligations are. That is, we should think that whereas having volitional control over whether one ϕ s is neither necessary nor sufficient for ϕ to be an option (and, thus, potentially an obligation), having rational control over whether one ϕ s is both necessary and sufficient for ϕ to be an option.

Keep in mind that this is a very tentative proposal, for I recognize that some will object to it. For instance, Fischer and Ravizza (1998) might object that by insisting that whether or not she ϕ s at t'' just depends on whether and how she responds at t' to her reasons, I'm wrongly presupposing that it's regulative control as opposed to guidance control that matters. But I'm willing to revise the above account so as to meet this objection should it prove to have merit. Indeed, I'm willing to revise the above account in any way that preserves the three guiding thoughts that I've spelled out above.

To see both that having volitional control is insufficient and that having rational control is necessary, suppose that we find that, when given the choice between eating one marshmallow now and eating two marshmallows fifteen minutes from now, children of a certain age invariably choose to eat the one marshmallow now. And further suppose that we find that although they have volitional control over whether they refrain from eating the one marshmallow now in that whether they do so just depends on whether they form the intention to do so, they lack the capacity to recognize and respond appropriately to the reason that they have for deferring their gratification (i.e., the fact that they would receive a greater reward later) and so invariably form the intention to eat the one marshmallow now. But if they genuinely cannot recognize and respond appropriately to the reason that they have for deferring their gratification, it seems that we would not hold them to be obligated to defer their gratification. And this suggests that it is not enough to have volitional control if one's volitions are not themselves under one's rational control.

Moreover, we should think both that having volitional control is not necessary and that having rational control is sufficient for having an option and, thus, a potential obligation. For this would explain why we take ourselves to have obligations to form various reasons-responsive attitudes (such as desires, beliefs, and intentions) despite our lacking volitional control over whether we do so. Thus, the view that it is rational and not volitional control that matters nicely carves out the range of things with respect to which we can have obligations. It explains why we have obligations with respect to such things as beliefs, intentions, and voluntary acts but not with respect to such things as involuntary acts, sensations of phantom pain, or feelings of hunger, dizziness, and tiredness. For whereas we have rational control over the former class of things, we lack rational control over the latter class of things. And this is because the mechanisms that produce those things in the former class are reasons-responsive, whereas the mechanisms that produce those things in the latter class are not. For instance, I may be aware that there is no reason for me to feel pain in my missing limb, but the mechanism that produces my sensations of phantom pain is not responsive to my awareness of this fact.

By contrast, the mechanism that produces beliefs in a rational being such as myself is responsive to my awareness of facts that constitute reasons for my having these beliefs.

Of course, some will object that we can't have obligations with respect to our beliefs, desires, intentions, and other reasons-responsive attitudes precisely because we don't have volitional control over them. But there are at least two problems with this view. First, it flies in the face of commonsense. We ordinarily presume that there are such obligations. We suppose that people are obligated to want what's best for their children, to believe what they have conclusive evidence for believing, and to intend the necessary means to their ends.²⁹ And we suppose this despite the fact that people lack volitional control over such attitudes.

Second, the idea that S could be obligated to ϕ only if S has volitional control over whether she ϕ s is unmotivated. It would be appropriate to infer "S is not obligated to ϕ " from "S does not have volitional control over whether or not she ϕ s" if we were to restrict the range of ' ϕ ' to (voluntary) actions. But the appropriateness of the inference when ' ϕ ' is restricted to actions has nothing to do with the putative necessity of having volitional control over them and everything to do with the nature of (voluntary) actions.³⁰ After all, what distinguishes an action (or, at least, a voluntary action) from a

²⁹ Technically, since I'm talking about only objective obligations, I should say that people are (objectively) obligated to believe what's true. People are only subjectively obligated to believe what they have conclusive evidence for believing. This is important given that I'll be arguing that which acts we're obligated to perform depends on which attitudes we're obligated to form. To illustrate, suppose that my evidence misleadingly suggests that I must push a button to save ten lives when in fact my pushing the button would destroy those ten lives. In this case, I claim that although I am subjectively obligated both to push the button and to believe that doing so would save the ten, I'm objectively obligated both to refrain from pushing and to believe that my pushing would destroy the ten.

Note also that we think that people are obligated to form (or not to form) certain beliefs and not merely to perform (or not to perform) the acts that would cause them to form (or not to form) these beliefs. To illustrate, imagine that I've promised to read an important book and so ought to read it. Further suppose that this book, although otherwise meritorious, contains some fallacious reasoning for the proposition p . Lastly, assume that I have no evidence for p . Thus, I ought not to form the belief that p . But, now, suppose that I am persuaded by this bit of fallacious reasoning and so form the belief that p as a result of reading this book. Here, then, is a case where I am criticizable for having formed the belief that p , and yet we cannot trace back the source of my responsibility to my having violated an obligation not to perform the act that led me to form this belief. After all, I was obligated to read the book. And I'll just stipulate that there was no other past act that I was obligated to perform that would have made me less disposed to being persuaded by this fallacious reasoning. For further reasons to reject this sort of tracing account of our responsibility for our reasons-responsive attitudes, see Smith 2015.

³⁰ This idea is borrowed from Chuard and Southwood 2009. For other similar responses to this sort of worry, see Graham 2014 (p. 400, n. 22), Hieronymi 2008, and Smith 2005 (pp. 264–265).

mere happening is that only the former is the type of event that's brought about at will and, thus, as a result of the agent's exercising her volitional control over that event.³¹ My foot rising as a result of my intention (or will) to kick a football is an action, whereas my foot rising as a result of a doctor's striking my patellar ligament with her reflex hammer is a mere happening. And this is because the former, but not the latter, was caused by my intention (or will). So if we restrict ' ϕ ' to actions, then S will necessarily have volitional control over her ϕ -ing simply because her actions are just those events over which she has volitional control. But this leaves us with no reason to think that, more generally, S can't be obligated to ϕ unless S has volitional control over whether or not she ϕ s.

Others may object that S can't be obligated to ϕ unless ϕ is the sort of thing for which she can be appropriately held responsible and that she cannot be appropriately held responsible for having ϕ ed unless she had volitional control over whether or not she ϕ ed. But although it is plausible to assume that S must have had some sort of control over whether or not she ϕ ed for it to be appropriate to hold her responsible for having ϕ ed, I don't see any reason for thinking that the relevant sort of control must be volitional control as opposed to rational control. After all, we often hold people responsible for their beliefs, desires, and intentions, and this seems entirely appropriate despite the fact that these attitudes are not under their volitional control. As Scanlon notes, "Because 'being responsible' is mainly a matter of the appropriateness of demanding reasons, it is enough that the attitude in question be...one that either directly reflects the agent's judgment or is supposed to be governed by it. For this reason, one can be responsible not only for one's actions but also for intentions, beliefs, and other attitudes" (Scanlon 1998, 22).

So I think that we should accept rationalism as opposed to voluntarism. And this suggests that we should combine possibilism with rationalism as opposed to voluntarism, thereby arriving at

³¹ Here, I'm assuming the causal theory of action, which stems from the influential work of Donald Davidson—see, for instance, his 1963.

Rationalist Possibilism: For any time t , subject S , and maximal option μ for S , S ought, as of t , to μ if and only if, and because, S 's μ -ing has feature F . For any time t , subject S , and non-maximal option ν for S , S ought, as of t , to ν if and only if, and because, there is a maximal option μ for S such that S ought, as of t , to μ and S 's μ -ing logically necessitates S 's ν -ing. And, for any ϕ , subject S , and time t , S has, as of t , the option of ϕ -ing if and only if, and because, S has, as of t , rational control over whether or not she ϕ s.³²

This is a form of possibilism because it distinguishes between maximal and non-maximal options and treats the deontic statuses of the latter as derivative of those of the former. On this view, we first evaluate maximal options according to whether or not they have feature F and then we evaluate a given non-maximal option according to whether or not its performance is logically necessitated by the performance of any maximal option that has feature F .

Now, when I first introduced possibilism, I had assumed that we were to evaluate maximal options by comparing their prospects and our reasons for preferring some of them to others. And, given the ecumenical nature of this approach, this was meant to be neutral between various substantive accounts of how maximal options are to be evaluated. But now that we realize that maximal options will include attitudes as well as actions, this approach no longer seems sufficiently ecumenical. After all, it is quite controversial to suppose that we are to evaluate attitudes in terms of their prospects. Many would, for instance, deny that I ought to believe that I have an immortal soul so long as

³² This view draws some inspiration from Holly Goldman's G^* principles—see her 1978. Her G^* principles are used to evaluate decisions as well as actions, and she thinks that it is important to evaluate decisions as well because, as she points out, our decisions are to some extent under our voluntary control, just as our actions are. But my view is importantly different from hers. For one, I take the relevant sort of control to be rational control, not volitional/voluntary control. For another, I think that we must include not only decisions (or intentions) in our evaluations, but also all other reasons-responsive attitudes, including beliefs and desires.

The view draws even more inspiration from Ellie Mason's view. She writes: "I do affect the world, and a moral theory should guide me in so far as I have choices about how I affect the world—consequentialism (a particularly invasive moral theory) claims jurisdiction over any and all utility production I am responsible for. My contribution to the world is through my entire life – through every aspect of my life that I can control. This is the crucial difference between acts or motives and the entire set of controllable elements of my life—my acts or my motives are only a part of my contribution to the world, and so there is no reason at all to think about them in isolation from the context in which they are productive of utility" (2002, 301). My view differs from hers, though, in that I don't think that we should evaluate every controllable element in terms of its utility production. That believing that p would produce a lot of utility is not a reason to believe that p , but only a reason to want to believe that p and to act so as to cause oneself to believe that p .

the prospect of my having this belief is preferable to the prospect of my not having this belief. Many would instead argue that whether I ought to have this belief depends on whether it is true—or, at least, on whether I have sufficient evidence for its truth. Consequently, I have, in the above formulation of rationalist possibilism, replaced talk of prospects with a placeholder—namely, F—for whatever the correct criterion for evaluating maximal options happens to be. And since these maximal options include everything over which we have rational control, including both actions and attitudes, we should think that ‘F’ won’t stand for anything so simple as ‘maximizes utility’. Rather, we should think that it stands for something considerably more complex, such as ‘contains only beliefs that are true, only desires for what’s valuable, and only intentions to perform acts with excellent prospects’. (I don’t think that even this is sufficiently complex and sophisticated, but it’s just an example.)

In any case, that agents ought to perform all and only those maximal options that have feature F is meant to be trivially true given that ‘F’ is just a placeholder for whatever the correct criterion is. Of course, this means that I can’t appeal to rationalist possibilism, as it’s formulated above, to arrive at any substantive verdicts about maximal options. But, in what follows, I will rely, not on the above formulation, but on our considered intuitions to arrive at verdicts about maximal options. This should be unproblematic so long as I consider, as I intend to, only uncontroversial cases. For if the cases are sufficiently uncontroversial, then, whatever the correct account of F is, it must be able to account for such intuitions.

We can get a feel for rationalist possibilism by considering its implications in the above cases, starting with *Two Drugs*. With regard to this case, it’s clear that any maximal option that Dr. Singh ought, as of t_0 , to perform will include her forming at t_1 the intention to give Patrick *both* A and B at t_2 and then giving him both A and B at t_2 . And any such maximal option will logically necessitate her giving Patrick A at t_2 . Thus, rationalist possibilism implies that Dr. Singh ought, as of t_0 , to give Patrick A at t_2 . And it also implies that Dr. Singh ought, as of t_0 , to give Patrick both A and B at t_2 . Likewise, in *Professor Malice*, rationalist possibilism implies both that he ought, as of t_0 , to accept the

invitation to review the book and that he ought, as of t_0 , to both accept the invitation and write a fair review. So, unlike actualism, rationalist possibilism is compatible with both INH and CIR.

With regard to *The Safe*, rationalist possibilism denies that Safar is, as of t_0 , obligated to enter the correct combination. For, according to rationalism, entering the correct combination is not even an option for him; it is not, after all, under his rational control. Even if he were to respond appropriately at t_1 to all his reasons, forming at t_1 the intention to try to enter what he has just been told is the correct combination (viz., 8675309), he would not enter 8675309, for he is incapable of holding more than one single-digit number in his mind at any given time, and he is not so lucky as to just guess the correct single-digit number at each point in time. So, unlike schedulist possibilism, rationalist possibilism doesn't have an overly broad conception of what our options are.

And, with regard to *Professor Procrastinate* (even the version in which 'Pre' and 'Post' both refer to one and the same person—viz., Professor Procrastinate), rationalist possibilism implies neither that Professor Procrastinate is, as of t_0 , obligated to accept the invitation at t_1 nor that he is, as of t_0 , obligated both to accept at t_1 and to start work at t_{100} . For starting work at t_{100} is not, on rationalism, an option for Procrastinate as of t_0 . This is not under his rational control, for no matter how he responds at t_1 to his reasons, he is not going to start work on the review when the book arrives at t_{100} . And given this and the fact that he's never going finish the review if he doesn't start working on it at t_{100} , he ought, as of t_0 , to decline the invitation at t_1 so that the journal will get the next best person to write the review. Thus, unlike schedulist possibilism, rationalist possibilism is concerned not with whether Pre and Post are numerically identical but only with whether Pre can effect Post's writing the review, which he can't.

So it seems to me that rationalist possibilism gets intuitive verdicts in all the above cases. Of course, voluntarist possibilism gets intuitive verdicts in all these cases as well. But I've argued that voluntarist possibilism has too narrow a conception of options, for it denies that we have the option of involuntarily forming various reasons-responsive attitudes. And, as we'll see in the next section, this leads to voluntarist possibilism's hav-

ing some counterintuitive implications in other cases, cases where rationalist possibilism gets the right answer. So, in the next section, I'll consider two more cases, cases where alternatives to rationalist possibilism (including voluntarist possibilism) get into trouble. And, in the process, we'll be able to explore two important features of rationalist possibilism.

3. *The Implications of Rationalist Possibilism: Two Buttons and The Incapacitation*

Consider a modified version of a case that comes from Donald Regan (1980). It's an interesting case and is particularly important in motivating rationalist possibilism, because it shows that an adequate practical theory will not be *exclusively act-orientated* but will instead be *attitude-dependent*, as rationalist possibilism is. To say that a theory is not exclusively act-orientated is to say that it requires something more of agents than just the performance of certain voluntary acts.³³ Rationalist possibilism is not exclusively act-orientated in that it requires agents to involuntarily form certain attitudes as well as to perform certain voluntary acts. And to say that a theory is attitude-dependent is to say that it, like rationalist possibilism, holds that which acts agents ought to perform depends on which attitudes they ought to form.³⁴

I'll call my modified version of Regan's case *Two Buttons*, because it involves two individuals, Coop and Uncoop, each with a button in front of him. It's now t_0 and depending on whether each pushes his button at t_4 , the consequences will be as depicted in Table 3. Assume that the resulting utilities will be evenly distributed over all parties, including both Coop and Uncoop.

Coop (the cooperative one) will desire at t_3 that they both push. But, unfortunately, Uncoop (the uncooperative one) will lack this desire at t_3 . And Coop will push his button at t_4 if, and *only if*, he finds at t_3 that Uncoop desires that they both push. Now, I'll just stipulate that they will each be able to read the other's mind at t_3 and thereby come

³³ 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 had in mind.

³⁴ In not being exclusively act-orientated and instead being attitude-dependent, rationalist possibilism differs from my earlier view, viz., securitism (see Portmore 2011, 177).

to know what the other desires at t_3 .³⁵ So Uncoop will not push at t_4 , for he has no interest in pushing. And Coop will not push at t_4 because he will read Uncoop's mind at t_3 and find that he lacks the desire that they both push. So, neither Coop nor Uncoop is going to push at t_4 , but we can still ask of each of them: "Which act, push or not-push, ought he to perform at t_4 ?"

Table 3

<i>Two Buttons</i>	Uncoop: Push at t_4	Uncoop: Not-push at t_4
Coop: Push at t_4	100 utiles	0 utiles
Coop: Not-push at t_4	0 utiles	10 utiles

According to actualism (whether binary or multitudinous), the answer is surprisingly that each of them ought, as of t_0 , to not-push at t_4 . Actualism implies that Coop ought, as of t_0 , to not-push at t_4 , because, given that Uncoop is going to not-push at t_4 , the prospect of Coop's pushing at t_4 is a 100% chance of 0 utiles, whereas the prospect of Coop's not-pushing at t_4 is a 100% of 10 utiles. And Coop should, we'll assume, prefer the prospect with the greatest expected utility. Moreover, actualism implies that Uncoop ought, as of t_0 , to not-push at t_4 . For given that Uncoop won't desire at t_3 that they both push, Coop isn't going to push at t_4 . And given that Coop isn't going to push at t_4 , the prospect of Uncoop's pushing at t_4 is a 100% chance of 0 utiles, whereas the prospect of Uncoop's not-pushing at t_4 is a 100% chance of 10 utiles. And Uncoop should, we'll assume, prefer the prospect with the greatest expected utility.

Thus, they're each to hold fixed the fact that the other is going to not-push and conclude, therefore, that they are to not-push themselves. Yet whether Coop pushes is under Uncoop's rational control. If Uncoop were to respond appropriately to his reasons and form the desire that they both push, Coop would, then, push. The problem with ac-

³⁵ If you don't like fanciful examples involving mind readers, then just imagine that Coop and Uncoop are both poker players, that Uncoop has an involuntary tell with respect to his desires (one that he cannot fake), and that Coop is able to read Uncoop's tell. And if you don't think that there are rational requirements with respect to desires, then just change the example so that whether Coop pushes depends on whether Uncoop believes something that he is required to believe but doesn't believe.

tualism, then, is that it is exclusively act-orientated and so cannot require Uncoop to “do” all that is necessary to ensure the optimal outcome. To ensure the optimal outcome, Uncoop must do more than just perform the voluntary act of pushing. Uncoop must additionally form the desire that they both push, and forming this desire is not a voluntary act. Indeed, the only way that Uncoop can come to desire that they both push is by recognizing and responding appropriately to the decisive reason he has for forming this desire—specifically, the fact that their both pushing would produce what’s best for them both.³⁶ Thus, being exclusively act-oriented, actualism is unable to require Uncoop to involuntarily form the desire that they both push in response to the decisive reason he has for doing so. (Voluntarist possibilism, which is also exclusively act-oriented, has this problem as well.)

Of course, there may seem to be an easy and obvious fix: simply conjoin actualism with the view that agents ought to form all the attitudes that they have decisive reason to form and only those attitudes that they have sufficient reason to form.³⁷ Call this *conjunctive actualism*, for it conjoins actualism with the view that one should form reasons-responsive attitudes in accordance with one’s reasons for doing so. But conjunctive actualism won’t do, for although it requires Uncoop to form the desire that they both push, it doesn’t require Uncoop to push. Given that Uncoop is not going to form the obligatory desire, conjunctive actualism still implies that Uncoop should not-push. What we need, then, is a theory that requires that Uncoop both forms prior to t_3 the desire that they both push and pushes at t_4 . For only a theory that requires both of Uncoop will require Uncoop to “do” all that is necessary to secure the optimal outcome.

Nevertheless, as I will show presently, even this is not enough. That is, it isn’t enough for a theory to require Uncoop both to form by t_3 the desire that they both push and to push at t_4 . An adequate practical theory must be attitude-dependent. That is, it

³⁶ Note that there’s no wrong-kind-of-reason problem here. The reason that Uncoop has to desire that they both push is not that his being in the state of desiring that they both push would be good (though it would be), but that the object of this desire—their both pushing—is good.

³⁷ 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 .

must hold that what an agent ought to do depends on what attitudes she ought to form. To see why, consider the following view.

Cooperativism: A subject ought to perform x if and only if she must perform x in order to do her part in the best cooperative scheme. More precisely, for any time t , subject S , and act x that is an option for S , S ought, as of t , to perform x at t' if and only if, and because, S 's x -ing at t' is part of the best cooperative scheme (i.e., the best scheme of action involving all agents).

This view implies that Uncoop should push, for Uncoop's pushing is part of the best cooperative scheme: the one in which Coop and Uncoop both push. So if we were to conjoin the cooperative view with the view that agent's ought to form all the attitudes that they have decisive reason to form and only those attitudes that they have sufficient reason to form (and call this combined view *conjunctive cooperativism*), we would end up with a theory that requires Uncoop both to form by t_3 the desire that they both push and to push at t_4 . Thus, we would end up with a theory that requires Uncoop to "do" all that is necessary to secure the optimal outcome. However, even this is not enough. For such a theory would have counterintuitive implications in the following variant on *Two Buttons*. Call it *The Incapacitation*.

In this case, everything is as it is in *Two Buttons* except that a demon rewires Uncoop's brain, rendering him incapable of forming the desire that they both push. Thus, in this case, Uncoop does not have the option of forming the desire and so cannot be obligated to do so. Yet conjunctive cooperativism still requires Uncoop to push, because his pushing is still part of the best cooperative scheme: the one in which he and Coop both push. But, given that Uncoop's pushing without desiring that they both push would have terrible consequences, we should think that Uncoop is required to push only if he is also required to form the desire that they both push. The problem with conjunctive cooperativism is that it requires Uncoop to push even if he is not required to desire that they both push. What we need instead is an attitude-dependent view, a view that makes what Uncoop is required to do dependent on which attitudes he is required to form.

So we've looked at four alternatives to rationalist possibilism: (1) actualism, (2) cooperativism, (3) conjunctive actualism, and (4) conjunctive cooperativism. The problem with first two is that they are exclusively act-orientated and so can't require Uncoop to desire that they both push—voluntarist possibilism has this problem as well.³⁸ The problem with conjunctive actualism is that although it requires Uncoop to desire that they both push, it fails to require Uncoop to push. And the problem with conjunctive cooperativism is that it is not attitude-dependent and, thus, holds that Uncoop should push even if he is not required to desire that they both push.

Rationalist possibilism is superior to these alternatives. First, not being exclusively act-orientated, it requires Uncoop, in *Two Buttons*, not only to push at t_4 but also to desire at t_3 that they both push. For Uncoop ought, as of t_0 , to perform a maximal option that includes (1) his forming at t_3 both the intention to push and the desire that they both push and (2) his pushing at t_4 . Second, being attitude-dependent, rationalist possibilism requires Uncoop to push at t_4 only if he ought to desire at t_3 that they both push. So, in *The Incapacitation*, rationalist possibilism implies that Uncoop ought, as of t_0 , to not-push at t_4 given that all the maximal options that Uncoop ought, as of t_0 , to perform will include his not-pushing and exclude his desiring that they both push. It will exclude his desiring that they both push, because this is not even an option for him given the demon's rewiring of his brain. And it will include his not-pushing, because this is what it would be best for him to do given that his lacking the desire that they both push is unalterable by him. So, in both *Two Buttons* and *The Incapacitation*, rationalist possibilism requires Uncoop to "do" all that is necessary to secure the best possible outcome. It's just that whereas, in *Two Buttons*, this is the outcome in which they both push, in *The Incapacitation*, this is the outcome in which they both not-push.

It seems to me, then, that rationalist possibilism gets the intuitively correct results in both *Two Buttons* and *The Incapacitation*. Yet some may question the plausibility

³⁸ Voluntarist possibilism implies that Uncoop should not-push in the *Two Buttons*. On voluntarist possibilism, maximal options consist only of acts and so do not include the involuntary formations of attitudes. So all the best maximal options available to him will include his not-pushing, for his not-pushing is what's best given that he's not going to form (in response to his reasons) the desire that they both push.

of rationalist possibilism's implications in *Two Buttons* once they realize that it implies not only that Uncoop is, as of t_0 , obligated to push at t_4 but also that Coop is, as of t_0 , obligated to not-push at t_4 . For it may seem absurd to hold both that Uncoop should push at t_4 and that Coop should not-push at t_4 given that the worse possible outcome (the one with 0 utiles) will obtain if Uncoop pushes at t_4 and Coop not-pushes at t_4 . But this is not absurd, for Uncoop and Coop are in very different situations. Uncoop is in a position to affect what Coop does, whereas Coop is not in a position to affect what Uncoop does. If Uncoop desires at t_3 that they both push (as he ought to), then Coop will push at t_4 . And this is why Uncoop is required both to have this desire and to push. However, Uncoop is going to not-push regardless what Coop does or desires. And this is why Coop must just take it for granted that Uncoop will not-push and make the best of a bad situation by not-pushing as well.³⁹

Of course, some people could, despite all these arguments, maintain that Uncoop should not-push given that Coop is going to not-push, and they may even acknowledge that Coop's not-pushing is contingent upon Uncoop's failing to fulfill his obligation to form the desire that they both push.⁴⁰ But if they do, they must acknowledge that their position implies that Uncoop faces a normative dilemma for which there is no way out. For if Uncoop forms the set of the attitudes that he ought to form (the set that includes both a desire that they both push and an intention to push), he will push. And yet these same people maintain that Uncoop should not-push. To my mind, this is too great a cost. We should not accept that normative dilemmas (for which there is no way out) can arise without their having been some past normative failure. It seems that it should be possible for someone who has made no previous mistake to ensure that she does not end up

³⁹ If you still have any doubts, then just imagine the following variation of *Two Buttons*: *Two Buttons 2*. This case is exactly like *Two Buttons* except that in this case whether Coop will push depends, not on whether Uncoop desires that they both push, but on whether Uncoop tells Coop that he's willing to push. And in *Two Buttons 2* Uncoop is not going to tell Coop that he's willing to push. In this case, it seems that Uncoop should push for he should both tell Coop at t_3 that he's willing to push and push at t_4 . But Coop should not push, for Uncoop is not going to push and nothing he can do will change that.

⁴⁰ Also, those who deny that Uncoop ought to push at t_4 will have to deny the validity of deontic detachment, which holds that we are to infer that S ought to ϕ from both the fact she ought to ψ and the fact that she ought to ϕ if she's going to ψ (Greenspan 1975). After all, it ought to be that Uncoop desires at t_3 that they both push, and he ought to push at t_4 if he's going to desire at t_3 that they both push. Thus, from deontic detachment, it follows that Uncoop ought to push at t_4 .

criticizable. Yet, according to these people, Uncoop will be criticizable either for failing to form the attitudes that he is required to form or for pushing, which is what he is guaranteed to do if he forms the attitudes that he is required to form. Yet, in *Two Buttons*, there needn't have been any past mistake on Uncoop's part. For we can just stipulate that Uncoop arose *ex nihilo* at t_0 with all and only the attitudes and dispositions that he is stipulated to have at t_0 , and this includes the lack of any desire at t_0 that they both push. Still, it would be a future mistake (one for which he would be criticizable) for him to fail to form the desire at t_1 that they both push in the face of his current awareness of decisive reasons for forming this desire.

To sum up, I've argued that we should accept both that, in *Two Buttons*, Uncoop ought, as of t_0 , to push at t_4 and that, in *The Incapacitation*, Uncoop ought, as of t_0 , to not-push at t_4 . The only kind of practical theory that can get these two intuitive verdicts is one that, like rationalist possibilism, is not exclusively act-orientated but is instead attitude-dependent. This, I believe, constitutes good evidence in favor of rationalist possibilism and against views like actualism, cooperativism, schedulist possibilism, and voluntarist possibilism, which lack one or more of these two features.

4. Conclusion

I've argued for a radically different sort of practical theory, a theory that is not exclusively act-orientated but is instead attitude-dependent. According to such a view, we cannot understand what our practical obligations are merely by considering which acts we could perform and what features they have. We must also consider which attitudes we should form and which actions it would make most sense for us to perform if we were to form these attitudes. This means that most typical practical theories are mistaken. Theories such as rational egoism, Rossian deontology, act consequentialism, and rule consequentialism are mistaken because they are exclusively act-orientated and, thus, unable to require Uncoop to involuntarily form the desire that they both push in response to the decisive reason he has for doing so. And theories such as global consequentialism and virtue ethics are mistaken, for even though they are not exclusively act-orientated, they

fail to be attitude-dependent and thus fail to make which actions Uncoop ought to perform dependent on which attitudes he's ought to form.⁴¹ So one important implication of this paper is that most, if not all, of the theories currently on offer are inadequate. We should, therefore, turn our attention to developing a theory such as rationalist possibilism that meets both the constraint of not being exclusively act-orientated and the constraint of being attitude-dependent.⁴²

The only two theories that I know of that are not ruled out by these two constraints are Regan's (1980) cooperative utilitarianism and Adams's (1976) conscience utilitarianism.⁴³ Cooperative utilitarianism is untenable for reasons that I spell out in Portmore *forthcoming*. And conscience utilitarianism is, I believe, also untenable. Conscience utilitarianism—when understood as a theory about what we ought to do, all things considered—holds both that (1) “we have a...duty [all things considered] to do an act, if and only if it would be demanded of us by the most useful kind of conscience we could have” (p. 479) and that (2) we have a...duty [all things considered] to have the

⁴¹ Virtue ethics, as typically formulated, isn't attitude-dependent. Take Hursthouse's statement of its criterion of rightness: “An action is right iff it is what a virtuous agent would characteristically (i.e. acting in character) do in the circumstances” (2000, 28). On this view, whether Uncoop should, in *Two Buttons*, push or not-push just depends on which of those two acts a virtuous agent would perform in his circumstances, not on which attitudes he (the actual agent) should form, nor on how those circumstances would change if he did. Given that Uncoop's circumstances are such that Coop is going to not-push, it seems that a virtuous person would also not-push so as to bring about the most utility possible in the circumstances. So, unlike rationalist possibilism and other attitude-dependent theories, virtue ethics implies that Uncoop ought to not-push. Attitude-dependent theories yield a different verdict because they're concerned, not with what it would be best to do in the actual circumstances, but with what it would be best to do in the circumstances that would be actual if the agent were to form all the attitudes that she ought to form. Of course, virtue ethics could be revised so as to be attitude-dependent. We could, for instance, reformulate its criterion of rightness along the following lines: *S*'s ϕ -ing is permissible if and only if ϕ is what a virtuous agent would characteristically do in the circumstances that would be actual if the agent were to form all the attitudes that she ought to form. If Uncoop were to form all the attitudes that he ought to form, his circumstances would be such that Coop is going to push. And, in those circumstances, a virtuous agent would also push.

For present purposes, I'm construing theories like virtue ethics, Rossian deontology, and act-consequentialism to be theories not only about what agents morally ought to do, but also about what agents ought to do, all things considered.

⁴² We can, of course, form rationalist-possibilist analogues of the sorts of theories that are currently on offer by adopting the relevant views about what feature *F* is. For instance, on utilitarian rationalist possibilism, we just substitute ‘maximizes expected utility’ for ‘has feature *F*’ in the formulation of rationalist possibilism. And, on Kantian rationalist possibilism, we just substitute ‘is in accordance with the Categorical Imperative’ for ‘has feature *F*’.

⁴³ Strictly speaking, resolute-choice theories such as those endorsed by Bratman (1999), Gauthier (1994), and McClellan (1990) are attitude-dependent, for they hold that whether a subject ought to perform *x* at *t*’ can depend on whether she had rightly planned/intended at *t*’ on not performing *x* at *t*’. Such theories, however, do not have the resources to deal with cases such as *Two Buttons*, where what Uncoop ought to do doesn't depend on whether he had previously formed the intention to push, but on whether he will form the desire that they both push.

most useful kind of conscience we could have. Note that, by ‘most useful’, Adams means ‘most productive of utility’. One reason to think that conscience utilitarianism is untenable is that it requires agents to have attitudes that they are clearly not rationally permissible. For instance, if the belief that God punishes those who fail to maximize utility with eternal damnation would be part of the most useful kind of conscience that I could have (because, say, it would motivate me to maximize utility), then conscience utilitarianism requires me to have this belief even if it’s false and even if I have no evidence for it.

In any case, I’ve shown more than just that rationalist possibilism can meet the above two constraints. I’ve also shown that it gets plausible verdicts in a wide-range of cases, including *Two Drugs*, *Professor Malice*, *Professor Procrastinate*, *Two Buttons*, and *The Incapacitation*. Lastly, I’ve shown that rationalist possibilism has important implications with respect to what counts as an option for an agent, implying both that agents can have the option of involuntarily forming certain attitudes and that agents can’t have the option of performing a future act if it’s one that she wouldn’t perform no matter how she responds at present to her reasons. Thus, it isn’t enough that she would perform that future act if she were to have the right intentions at the right times, she must be able to secure her performance of that future act by responding appropriately to her reasons at present.⁴⁴

Arizona State University
dwpportmore@gmail.com

References

Adams, R. M. (1976). “Motive Utilitarianism.” *Journal of Philosophy* 73: 467–481.

⁴⁴ For helpful comments and discussions on precursors to this paper, I thank audiences at Syracuse University, Arizona State 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 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. Shaha, 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.

- Bratman, M. (1999). "Toxin, Temptation, and the Stability of Intention." In his *Faces of Intention*, pp. 58–90. Cambridge: Cambridge 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.
- Cariani, F. (2013). "'Ought' and Resolution Semantics." *Noûs* 47: 534–558.
- Chuard, P., and N. Southwood (2009). "Epistemic Norms without Voluntary Control." *Noûs* 43: 599–632.
- Davidson, D. (1963). "Actions, Reasons, and Causes," *Journal of Philosophy* 60: 685–700.
- Feldman, F. (1986). *Doing the Best We Can: An Essay in Informal Deontic Logic*. Dordrecht: D. Reidel Publishing Company.
- Fischer, J. M., and M. Ravizza. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Gauthier, D. (1994). "Assure and Threaten." *Ethics* 104: 690–716.
- Graham, A. (2014). "A Sketch of a Theory of Moral Blameworthiness." *Philosophy and Phenomenological Research* 88: 388–409.
- Greenspan, S. (1975). "Conditional Oughts and Hypothetical Imperatives." *Journal of Philosophy* 72: 259–276.
- 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.
- Hare, C. (2011). "Obligation and Regret When There is No Fact of the Matter About What Would Have Happened if You Had not Done What You Did." *Noûs* 45: 190–206.
- Hieronymi, P. (2008). "Responsibility for Believing." *Synthese* 161: 357–373
- Howard-Snyder, F. (1997). "The Rejection of Objective Consequentialism." *Utilitas* 9: 241–248.
- Hursthouse, R. (2000). *On Virtue Ethics*. Oxford: Oxford University Press.
- Jackson, F. (1988). "Understanding the Logic of Obligation." *Proceedings of the Aristotelian Society*, Supplementary Volumes, 62: 255–283.

- . (2014). "Procrastinate Revisited." *Pacific Philosophical Quarterly* 95: 634–647.
- 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.
- Mason, E. (2002). "Against Blameless Wrongdoing." *Ethical Theory and Moral Practice* 5: 287–303.
- McClennen, E. (1990). *Rationality and Dynamic Choice*. Cambridge: Cambridge University Press.
- McNamara, P. (2014). "Deontic Logic." In E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), URL = <<http://plato.stanford.edu/archives/spr2014/entries/logic-deontic/>>.
- Parfit, D. (2011). *On What Matters*. Vol. 2. Oxford: Oxford University Press.
- Portmore, D. W. (2011). *Commonsense Consequentialism: Wherein morality meets rationality*. New York: Oxford University Press.
- . (forthcoming). "Consequentialism and Coordination: How Consequentialism Has an Attitude Problem." In C. Seidel (ed.), *Consequentialism: New Directions, New Problems?* Oxford: 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.
- Sinnott-Armstrong, W. (1984). "'Ought' Conversationally Implies 'Can'." *Philosophical Review* 93: 249–61.
- 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>.
- . (2010). "Attitudes and Control." Unpublished manuscript dated April 12, 2010.
- . (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115: 236–271.

- Streumer, B. (2003). "Does 'Ought' Conversationally Implicate 'Can'?" *European Journal of Philosophy* 11: 219–228.
- Vessel, J-P. (2003). "Counterfactuals for Consequentialists." *Philosophical Studies* 112: 103–125.
- von Wright, G. H. (1956). "A Note on Deontic Logic and Derived Obligation." *Mind* 65: 507–509.
- Wedgwood, R. (2006). "The Meaning of 'Ought'." In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Volume 1, pp. 127–160. Oxford: Oxford University Press.
- Wiland, E. (2005). "Monkeys, Typewriters, and Objective Consequentialism." *Ratio* 18: 252–360.
- Woodard, C. (2009). "What's Wrong with Possibilism." *Analysis* 69: 219–226.
- Zimmerman, M. J. (1996). *The Concept of Moral Obligation*. Cambridge: Cambridge University Press.