THE BEST ARGUMENT FOR “OUGHT IMPLIES CAN” IS A BETTER ARGUMENT AGAINST “OUGHT IMPLIES CAN”

BRIAN TALBOT

Washington University in St. Louis

To argue that “ought” implies “can,” one can appeal to general principles or to intuitions about specific cases. One general truism that seems to show that “ought” implies “can” is that obligations must be able to guide action, and putative obligations that are unfulfillable are unable to do so. This paper argues that obligations that are unfulfillable can still guide action, and that moral theories which reject the principle that “ought” implies “can” are actually better able to account for how obligations guide than theories which endorse “ought” implies “can.” The paper also argues that any intuitions about specific cases that seem to provide evidence that “ought” implies “can” do not actually give us this evidence. Rather, these intuitions pose similar problems for theories which accept “ought” implies “can” as they pose for theories which reject the principle. Some theories which reject “ought” implies “can” will fit our intuitions at least as well, if not better, than theories which accept it. So, intuitions do not favor accepting that “ought” implies “can,” and appeal to general principles favors its denial.

1. Introduction

The principle that “ought” implies “can” (OIC) says, roughly, that if one ought to φ , then one must be able to φ .¹ While some see OIC as obviously true, there are too many apparent counterexamples to the principle to take its truth for granted; rather, it must be argued for. This paper will consider what seem to be the two best ways of arguing for OIC. It will show that they do not support the principle, but instead undermine its plausibility.

1. There are different versions of the principle—for example, some say that if one ought to φ at time T , then one must be able to φ at T , while others say that one must only have been able to φ at some previous time (contrast Zimmerman, 1996, with Howard-Snyder, 2006)—but these variations won’t matter for my paper.

One way of arguing for OIC is to appeal to general claims about the nature of oughts and obligations. It seems a truism that oughts give us guidance, that obligations must (at least in theory) be able to channel action in one direction or another. This provides an often-cited justification for OIC: when an agent cannot φ , a putative obligation to φ obviously cannot guide that agent to φ , and it seems that such an obligation would be pointless, inert.² Thus, the argument for OIC goes, a putative unfulfillable obligation cannot be a real obligation. It turns out, however, that accepting the truism that obligations must guide does not require accepting OIC. In Section 2 of this paper, we'll see how an obligation to φ can guide even when the obligated agent cannot φ . With this understanding of guidance in mind, we'll see in Section 3 that theories which deny OIC do a better job of accounting for how oughts guide than theories which accept OIC.

While the principle that oughts must be able to guide gives a relatively abstract, theoretical argument for OIC, many are moved to accept OIC by consideration of concrete (or possibly concrete) specific situations. Situations that support OIC must be ones in which an agent cannot φ and in which, intuitively, it is not the case that the agent ought to φ .³ OIC seems to give an easy explanation for the truth of these intuitions, one not available to theories which deny OIC. However, employing the notion of guidance developed in Sections 2 and 3, I will show in Section 4 that these intuitions actually pose similar problems for theories which accept OIC as they do for theories that deny OIC. Some theories which deny OIC are at least as capable of fitting these intuitions as theories that accept OIC, and potentially more capable. This means that no intuitions about cases provide support for OIC, and the intuitions which seemed to support OIC might instead push us to deny it.⁴

I focus on moral oughts in this paper. I will very briefly sketch in Section 5 how my argument might apply to moral reasons as well. What I say may be applicable to non-moral oughts, such as epistemic ones, but I won't have the

2. See, for example, Hare (1963), Williams (1995), and Copp (2003). Arguments in Vranas (2007) can be interpreted along these lines as well.

3. For arguments for OIC that use such intuitions, see Howard-Snyder (2006), Vranas (2007), and Lowry (2012). Streumer (2007) appeals to intuitions like this to argue that "reason to" implies "can," and in Streumer (in press) uses this to argue that "ought" implies "can." In discussing this paper with others, it is these sorts of intuitions that were most often cited as motivating acceptance of OIC.

4. There are some other ways of arguing for OIC that I cannot address in this paper. There is evidence for OIC from linguistic practices, although this evidence is also consistent with the claim that "ought" merely conversationally implies "can" (see Sinnott-Armstrong 1984). It is also extremely hard to give a good logic of obligations when we deny OIC (Howard-Snyder, 2006, argues for this; see Goble, 2009, for further discussion and some deontic logics without OIC). Finally, David Copp (2003) has argued that it would be unfair if OIC were false; see Footnote 23 for a brief discussion. In many cases, however, those who endorse OIC do not really argue for the principle, but focus on responding to objections (e.g., Zimmerman 1996; Vranas 2007).

space to explore this. This paper is intended to be relatively ecumenical, and I'll try to avoid arguments that rely on the truth of any specific moral theory.⁵ This is because arguments for OIC are typically made independently of any specific moral theory (see, e.g., Copp 2003; Howard-Snyder 2006; Vranas 2007), and it is often taken for granted that any plausible moral view must say that OIC is true (e.g., Zimmerman 1996).

Let's clarify some terms before getting started. As is standard in this literature, I will assume that "ought" and "obligation" are synonymous, and will use "one ought to φ " and "one is obligated to φ " interchangeably. I will also assume, as is standard, that when one is obligated to φ or ought to φ , φ is required and it is wrong or impermissible to not φ .⁶ I'll use the term "unfulfillable obligation" to refer to an agent's obligation to φ at some time when that agent is unable to φ , for the relevant sense of "unable." So, an agent has an unfulfillable obligation when the agent ought to φ but cannot φ . There are diverging views on what the "can" in OIC means. I will remain neutral on this topic. The overall arguments I give can be applied to all of the options. However, some of my examples may need adjustments to work for particular interpretations of "can." This is sometimes because these examples are only supposed to be relevant to certain understandings of "can." In other cases, filling in the details to make an example fit every possible interpretation of "can" would make the paper harder to read without adding much value.

2. Unfulfillable Obligations and Robust Action Guidance

Let's start by talking about different ways in which unfulfillable obligations might guide.⁷ The ideas we discuss in this section will be central to all the arguments in this paper. It will be easiest to explain these ideas by starting with some illustrative examples. It won't hurt my arguments if you don't agree with the claims I make here about any particular example, as long as some of what I say seems plausible in some relevantly similar situations.

5. My arguments will not convince utilitarians, but I suspect that this is because utilitarians really don't believe in obligations, at least not as non-utilitarians understand them. See Footnote 18.

6. To see this assumption made explicit in the discussion of oughts or OIC, see, e.g., Singer (1972), Zimmerman (1996), Ryan (2003), Howard-Snyder (2006), Vranas (2007), and Goble (2009). Those who reject the synonymy of "ought" and "obligatory" can read my paper as arguing that our evidence does not show that "obligation to φ " implies "can φ ." On this reading my thesis is still quite controversial.

7. The idea that unfulfillable obligations might guide in some or all of the ways I discuss in this section has been advanced by other authors, such as Brown (1977), Marcus (1980), Sinnott-Armstrong (1984), Heuer (2010), and Graham (2011).

Perhaps the most widely discussed unfulfillable obligations in the literature on OIC are obligations which an agent makes herself unable to fulfill. To adapt an example from Walter Sinnott-Armstrong (1984), imagine that Carrie makes a promise to pick Mindy up at noon. Carrie then gets on a flight at eleven, which makes her unable to fulfill her promise at noon. If she still ought at noon to keep her promise, this obligation could not guide her to pick Mindy up, since she just can't do so. However, the obligation might provide guidance in other ways. Carrie had reasons, at eleven, to not get on the plane. It is natural to say that these reasons arose from Carrie's future obligation to pick Mindy up. If that's right, then the now-unfulfillable obligation still has provided guidance in the past. The now-unfulfillable obligation can also provide guidance in the future. When Carrie fails to pick Mindy up, she owes Mindy an apology, and should make things up to her. It is natural to explain this by appealing to the wrongness of not picking Mindy up. Being wrong is equivalent to being a violation of an obligation. So Carrie's unfulfillable obligation might guide Carrie to apologize or atone.

Another sort of unfulfillable obligation is one which ordinary agents can fulfill, but which a particular agent cannot, due to facts about her nature or circumstances. For example, imagine that Laurel is a compulsive and psychopathic kleptomaniac who is both unable to refrain from stealing but who endorses her own actions (the example is adapted from Ryan, 2003). Laurel, like Carrie, should apologize and provide restitution for what she does. If this is explained by the wrongness of her thefts, then her unfulfillable obligation to not steal provides guidance. We can stipulate that Laurel's character was never within her control, so that, unlike Carrie, there isn't any feasible past act or omission with which to explain her reasons to apologize or provide restitution. This example also illustrates a further, vital way that unfulfillable obligations might provide guidance. Imagine that Laurel steals some clothes which don't fit her and gives them to her friend. Laurel's friend might have reasons to not accept the gift in order to avoid complicity in Laurel's theft. If the reasons to avoid complicity are explained by the wrongness of the theft, then we have another way that the obligation to not steal can guide even if unfulfillable. This sort of third-party guidance is quite important, because it shows how oughts can guide even if the subject of the obligation is, like Laurel, completely unmoved by any of their moral reasons.

There may be oughts which are neither fulfillable by the obligated agent nor by ordinary agents. Good examples come from consideration of ideals.⁸ Consider the ideal to love all people as one loves oneself, or the ideal to bring one's emotions and intellect into full accord. If these ideals were obligations, or generated obligations, then they would be unfulfillable by any normal person. And

8. Dilemmas—in which an agent must φ and must ψ but cannot do both—might also involve obligations which in some sense no agent can fulfil.

yet they could still guide. If agents had reasons to approximate these ideals, and these reasons were explained by the obligation to fully conform to the ideals, then these unfulfillable obligations would still guide.⁹

Let's generalize. Consider A's obligation to φ at time T. The most obvious way that this obligation can guide is by guiding A to φ at T. But as we've just seen, there are other ways the obligation can guide. It can guide A before T, by telling A how to plan. It can guide A after T, by telling A how to respond to her past wrongdoing. It can guide A at T, by telling A to do acts other than φ , such as approximating the obligatory act or taking "second best" options. It can guide agents other than A, by telling them how to respond to A's wrongdoing. In the examples so far, we talked about guidance to apologize, atone (provide restitution), avoid complicity, and approximate. There are many other potential responses to future, present, or past violations of obligations that I haven't discussed—mitigating wrongs or trying to prevent violations, for example. These are all forms of what I'll call *robust action guidance*. Moral theories provide robust action guidance when they tell agents how to respond to future, past, or present actions (their own actions or those of other agents). A theory can explain robust action guidance in a variety of ways, as we'll see in the next section. If robust action guidance is explained in part by obligations—if the wrongness of future, past, or present actions sometimes explains how we should respond to them—then unfulfillable obligations can still guide. Here some may balk, saying that robust action guidance is not what they had in mind when they said that obligations must be able to guide. I will respond to this in Section 3.4, using arguments and ideas I develop throughout Section 3. For now, let's assume that, if unfulfillable obligations can generate robust action guidance, then they can guide action.

3. The Benefits of Denying OIC

So, accepting the truism that all obligations must guide does not mean accepting OIC. In this section, we'll see that accepting this truism provides us with good reasons to *deny* OIC. We'll see this by comparing how theories which accept OIC and theories which reject OIC can each explain robust action guidance. Why is this comparison at issue? Much of what we say about the above examples, or examples like them, should be true whether or not OIC is true. That is, theories that accept OIC should still say that Carrie has reasons to not make herself unable to pick Mindy up, that Carrie and Laurel have reasons to apologize and make restitution, that Laurel's friends have reasons to not benefit from Laurel's theft,

9. This idea was inspired by Staffel (2015).

or that non-ideal agents have reasons to approximate unfulfillable ideals. More generally, all theories must give robust action guidance in response to at least some actions that are outside of the acting agent's control, whatever the theory says about OIC. Theories which accept OIC and those which reject it must differ in how this robust action guidance is explained, since theories that accept OIC cannot explain all robust action guidance by appealing to obligations. This, I will argue, raises problems for theories which accept OIC.

I will use the term "*DOER theory*" for a theory that *Denies OIC and Explains Robust action guidance by appeal to obligations*. Such a theory might say that Laurel the kleptomaniac must apologize for her thefts because they are wrong. I'll compare DOER theories to *non-DOER theories* which accept OIC. I want to see what difference denying OIC, by itself, makes to the quality of a theory. So, I want to focus on comparisons between the *closest* individual DOER and non-DOER theories—those that agree on as many prescriptions as possible. I'm going to assume for now that, for any DOER theory, we can give some non-DOER theory that issues exactly the same prescriptions in all the same cases, and *vice versa*, with the obvious exception being that only DOER theories will say we have unfulfillable obligations. Since we are accepting for now that unfulfillable obligations can guide, we won't (for now) count this one difference against DOER theories.¹⁰ (In Section 4, we will consider whether DOER and non-DOER theories really might issue problematically different prescriptions.) So, to compare the merits of the closest DOER and non-DOER theories, we can only consider the merits of how they explain the prescriptions they issue.

Let's use two examples to focus our discussion.

Ordinary racist: Duke unjustly and intentionally discriminates against members of other races. This discrimination is endorsed by Duke. He is able to act otherwise.

Uncontrollable racist: Duke unjustly and intentionally discriminates against members of other races. This discrimination is endorsed by Duke. But his actions are due to facts about his character which he cannot change, nor has ever been able to change

To make it clear that Duke's actions in *uncontrollable racist* are subject to moral evaluation, I wanted to make it clear that they are really actions, rather than something that simply happens to his body, so I have stipulated that his behavior is intentional and self-endorsed. I have also constructed *uncontrollable racist* so that Duke cannot do otherwise in whatever sense of "can" is relevant to OIC.

10. Cases like Carrie's and Laurel's are really evidence against OIC, since it is intuitive that they have unfulfillable obligations (see, e.g., Sinnott-Armstrong 1984; Ryan 2003). In cases where agents cannot ϕ , but robust action guidance is still generated, intuitions will not favor OIC.

I have made the example a bit extreme so that the issues I want to address are salient, but in part I chose this example because it is not too far from the ordinary. Recent work in psychology suggests that those raised in societies with a long history of racism will inevitably internalize and act on racist views (see, e.g., Gendler 2008; 2011 for discussion). Not everyone in these societies will endorse their own racist attitudes or the behavior this produces, and one can work to reduce or mitigate racist behavior. But most people will endorse at least some of their own racist behavior, no one seems to be able to control all of it, and the extent to which we are able to reduce or not endorse our behavior is likely not up to us, but is rather substantially affected by environmental factors outside our control. Thus, it is quite plausible that ordinary people will sometimes engage in and endorse racist behavior produced by aspects of their character that they have never been in a position to do anything about.

Both DOER and non-DOER theories should agree that Duke does something wrong in the *ordinary racist* case. They should also agree on a number of things about *both* cases. First, Duke owes an apology to those he has discriminated against. Second, he has good reasons to provide restitution to those he has discriminated against. Third, if he can somehow mitigate his discrimination, even if he can't eliminate it, he should do so—he should try to at least approximate non-discriminatory behavior. Fourth, third parties have reasons to avoid complicity in his discriminatory behavior. For example, others have reasons to avoid benefitting from Duke's racism even when that wouldn't exacerbate the behavior, and reasons to speak up against his actions even when this won't be listened to. Whatever moral theory we accept, DOER or non-DOER, should say all or most of these things about both cases. Each of these describes a form of robust action guidance, because they have to do with how Duke or others should respond to Duke's racist behavior: had Duke acted otherwise, there would be no reasons to apologize, atone, approximate, and avoid complicity.¹¹

It's very important for my arguments that this robust action guidance is partly independent of Duke's ability to control his racist behavior. To clearly see that this is so, consider Duke's reasons to atone and third parties' reasons to avoid complicity. If Duke had no reasons to atone to those he discriminates against in *uncontrollable racist*, then no one would, and people would have been unjustly discriminated against and yet be owed nothing in compensation. This is implausible, to put it mildly.¹² It should also be clear that third parties have reasons to avoid complicity in Duke's behavior even when Duke cannot control it. This is not just to say that they should not further the harm done; they should also avoid involvement in the discrimination even when that has no impact on

11. Note that I haven't yet talked about blame or punishment. See footnote 17.

12. I am not saying that Duke should be punished. Restitution is not punishment, as restitution does not seek to harm Duke; it seeks instead to compensate the victims of his action.

the outcomes. It is obvious that one should not endorse, participate in, or benefit from racist behaviors even when the racist can't help himself. We'll return to these points below.

While I endorse everything I've just said, the arguments to follow don't depend on the two *racist* cases specifically. All my arguments need is that there are some pairs of cases that are structurally similar to them. These will be pairs of cases in which the main difference between the two is that an agent can φ in one of the pair but not the other, and yet the agents have similar reasons to apologize or make reparations for failing to φ in both, the agents have similar reasons to approximate φ ing in both, and third parties have similar reasons to avoid complicity in the failure to φ in both. All moral theories—DOER and non-DOER alike—must account for the common robust action guidance in these pairs of cases. We can then compare how they do so to see whether we should endorse a DOER or non-DOER theory.

DOER theories provide a relatively simple account of this robust action guidance. It is wrong for Duke to discriminate in each case, and the violation of the obligation to not discriminate explains Duke's reasons to apologize, atone, and approximate, as well as third parties' reasons to avoid complicity. The wrongness of Duke's discrimination is due to some morally-relevant features of the act or situation, features shared between *ordinary* and *uncontrollable racist*. For example, the obligation to not discriminate might come from the harm discrimination does, the need to respect human dignity, and the connection this discrimination gives Duke to larger patterns of social injustice. These features are present in both *ordinary* and *uncontrollable racist*. More generally, for other parallel cases like the *racist* examples, DOER theories explain fulfillable and unfulfillable obligations by appeal to shared morally-relevant features, and explain shared robust action guidance by appeal to these obligations.

In the next two subsections, we'll consider the two possible forms non-DOER theories can take (if they accept OIC). Any specific theory which takes one of these forms has a significant disadvantage when compared to the closest DOER theory.

3.1. *Disjunctive Non-DOER Theories and the Doubling Up Problem*

There is something appealing about saying that reasons to apologize, atone, avoid complicity and so forth are usually explained by the wrongness of acts. A *disjunctive non-DOER theory* might try to hold onto this explanation as much as possible by giving a disjunctive explanation of robust action guidance. On such a view, when an agent has a fulfillable obligation to φ , this obligation explains robust action guidance. When the agent does not have a fulfillable obligation, robust action guidance comes from some other source.

To illustrate how disjunctive theories fare worse than DOER theories, it'll be helpful to consider a specific sort of disjunctive theory and a more fleshed-out version of the *racist* cases. The example disjunctive theory is based on ideas from Michael Zimmerman (1996), Frances Howard-Snyder (2006), and Bart Streumer (2007). On this view, when an agent can ϕ , this obligation generates robust action guidance. When the agent can't ϕ , any robust action guidance that exists is explained by past actions or omissions of the agent that were within the agent's control. To see how this theory might explain the robust action guidance in *uncontrollable racist*, we need to add more detail to the cases. Let's say that in both cases Duke is in a position of power in his workplace, and his racism manifests in the form of repeated and demeaning jokes made in front of minorities in his company who are in a position of lesser power and who depend on their jobs for survival. On this disjunctive view, Duke's past decision to take a job that put him in a position to treat people badly must be what explains the robust action guidance in *uncontrollable racist*—no other candidate action suggests itself, since his character has never been under his control.

Disjunctive non-DOER views like this run into what I call the *doubling up* problem. The doubling up problem is that, on the disjunctive non-DOER view, all of the robust action guidance seems implausibly weaker in *uncontrollable racist* than in *ordinary racist*. Let's imagine that in *ordinary racist* Duke knows himself well enough to predict that he will choose to act as he ultimately does, while in *uncontrollable racist*, Duke knows that he tends to behave in a racist manner but doesn't know himself well enough to know that this is uncontrollable. So, in both cases Duke's choice to put himself in a position of power is equally likely, as far as he can tell, to lead to his bad treatment of subordinates. If Duke's past choice to put himself in a position to discriminate by itself generated robust action guidance to atone, approximate, and avoid complicity in *uncontrollable racist*, then it should do the same in *ordinary racist*. But, on this disjunctive view, the wrongness of the actual discrimination in *ordinary racist* also gives guidance to atone, approximate, and avoid complicity. So there are two independent sources of guidance, which "doubles up" the robust action guidance. This suggests that the robust action guidance should be weaker in *uncontrollable racist* than *ordinary racist*, since it would only come from one source. I'm going to take a bit of time to explain why this is a problem. If you are already convinced that it is, skip to the next subsection (3.2) in which I discuss the issues with potential solutions to this problem.

If there were a difference in strength of robust action guidance between the two cases, it could be slight, or it could be fairly significant. If it were slight, it would have to be the common features of the two cases that contributed most of the strength to the robust action guidance. That is, most of the robust action guidance would be coming from Duke's past choice to put himself in a position

where he will act badly, and relatively little from Duke's wrong action in *ordinary racist*. If that is right, then this disjunctive view is really just a minor variation of a different non-DOER view—the elegant non-DOER view—whose very serious problems I'll discuss later (in 3.3). So let's instead consider why it would be problematic for the difference in strength of robust action guidance between *ordinary* and *uncontrollable racist* to be relatively significant.

To see why this would be implausible, we have to consider what it generally means for robust action guidance to be stronger or weaker. The stronger guidance is, the more able it is to override conflicting reasons: if Duke's robust action guidance gave him reasons to do one thing, but other moral considerations gave him reasons to do something else, we could determine what he should do by asking which reasons were stronger. On many views, stronger guidance can also be more demanding: if Duke had to make some personal sacrifice to conform to his robust action guidance, the weaker the robust action guidance was, the less of a sacrifice he would be required to make.

With that in mind, what might it mean for the strength of robust action guidance to vary significantly between *ordinary* and *uncontrollable racist*? Imagine that Duke's subordinates need therapy to recover from the treatment they receive at Duke's hands. Duke has reasons to pay for this therapy, and these reasons are a form of robust action guidance. If demandingness places constraints on obligations, then Duke would not have to pay for the therapy if it were sufficiently expensive. Were the reasons for him to atone for what he's done significantly weaker in *uncontrollable racist* than in *ordinary racist*, then the amount he would potentially be obligated to pay would be significantly lower. Similarly, we can imagine versions of the cases in which Duke has to break some other morally significant commitment to pay for this therapy—for example, he might have to break a promise to his daughter to buy her a pony for her birthday. Again, were the reasons for Duke to atone significantly weaker in *uncontrollable racist*, then they would be outweighed by much less important commitments than in *ordinary racist*. Given these sufficiently weaker reasons, Duke could truthfully say to his employees in some version of *uncontrollable racist*, "I have unjustly harmed you by actions of mine that I endorse. Ordinarily, I would be morally obligated to help you recover from what I have done to you. In fact, ordinarily a person could be obligated to pay much, much more than this (or to violate some other much more significant obligation in order to pay you). However, had I tried to not do these things to you I would have failed. And so I need not make up for what I have done, even though I did it on purpose." This is not plausible. From the perspective of the victims of Duke's action, little about what has been done to them changes between the scenarios, and so it makes little sense to say that Duke's obligations to make things up to those victims should change (I'll have something to say in a moment to those who do not share this intuition).

To bolster this point, let's consider another form of robust action guidance. Imagine that Duke's assistant Dwight is angling for a raise by ingratiating himself to Duke. To do so, he laughs at Duke's racist jokes, but only in private when the victims of those jokes are absent. (We make this last stipulation to ensure that Dwight's actions don't exacerbate the problem—they merely make him complicit in the behavior.) On both a DOER and disjunctive view, Dwight should receive robust action guidance to not act in this way. But his reasons to avoid complicity can be overridden, or might be overly demanding. If the robust action guidance in *ordinary racist* is much stronger than in *uncontrollable racist*, then a benefit to Dwight, or a commitment of his, that would never come close to justifying complicity in Duke's horrible behavior in *ordinary racist* can easily justify complicity in *uncontrollable racist*. As above, this is extremely implausible because the treatment that Dwight is a party to has not changed. If this implausible claim were true, then there would be some version of this case in which, upon finding out that Duke's behavior is uncontrollable, Dwight can truthfully exclaim, "Thank God that Duke can't help but treat people so badly, because *now* I'm permitted to laugh at his jokes!"

I've just said that the treatment Dwight is a party to, and to which Duke's employees have been subjected, does not change between the examples, and that the robust action guidance should not vary significantly in strength. Perhaps some will not share this latter intuition. This is because this issue is easily confounded with the question of whether those affected by Duke's actions would forgive him. It's plausible that, if Duke's subordinates knew that Duke could not control his actions, they would be likely to hold him less to account. They might even be less affected by his actions—it might be easier to ignore the uncontrollable racist. But this just tells us that, in some versions of these scenarios, the subordinates would either be merciful or be less harmed. Neither reflects a diminution of Duke's reasons to atone in the versions under discussion. To see why, imagine that in *uncontrollable racist* one of Duke's subordinates were to say, "I don't forgive you—pay to rectify the harm you have done to me!" Duke's response that he couldn't help being racist would not make a difference to the compensation the subordinate would be owed.

These examples illustrate a general issue. There will be plenty of parallel cases like *ordinary* and *uncontrollable racist* which share most of their morally-relevant features and differ just in whether an agent can control behavior which would ordinarily be wrong. On a disjunctive non-DOER view, the morally-relevant features in the uncontrolled case generate robust action guidance. These features will also be present in the parallel, controlled case. This suggests that the robust action guidance will be less strong in cases involving uncontrollable behavior than in cases involving controllable behavior. If the difference in strength is minor, then disjunctive non-DOER theories are really slight variants

of the non-DOER theories I'll discuss later. If the differences are significant, the disjunctive non-DOER theory issues implausible prescriptions. This is why the doubling up problem is a problem.

3.2. *The Problems with Solving the Doubling Up Problem*

Disjunctive theories can be given that deal with the doubling up problem. They can be made, like their closest DOER rivals, to give the right amount of action guidance in all the right places. The simplest way to solve the problem is just to stipulate that the source of robust action guidance that generates reasons when an agent cannot φ doesn't operate when an agent can φ . This would sacrifice some of the explanatory power of this theory, a sacrifice DOER theories don't make because they don't face the doubling up problem. That said, I'm confident that there are ways of fixing disjunctive theories that do not sacrifice explanatory power. However, any solution will necessarily add complexity to the fixed theory, complexity which the closest DOER alternative does not need because it does not face the doubling up problem.

Complexity is a general issue for disjunctive non-DOER theories.¹³ Even if we ignore the doubling up problem, disjunctive theories will be more complex than their closest DOER rivals. I should be clear what I mean by this.¹⁴ Theories can be simpler by being more parsimonious—by containing fewer or less complex types of entities. DOER and disjunctive non-DOER theories seem to contain all the same types of entities (reasons and obligations), and so neither is more parsimonious. Another sort of simplicity is elegance. While elegance tends to be less well defined than parsimony, it has to do with the number and complexity of hypotheses in a theory. The importance of elegance is apparently easier to defend than the importance of parsimony (Baker 2013). DOER theories are more elegant than their closest disjunctive rivals.¹⁵ DOER theories explain robust action guidance simply by appeal to obligations, whereas disjunctive theories necessarily appeal to obligations and some other source. If a disjunctive theory gives a satisfactory answer to the doubling up problem, even more complexity is needed.

A final problem with disjunctive theories is what I call *disunity*. To understand what disunity is, consider again the *racist* cases. If OIC is true, then in

13. See Graham (2011) for a discussion of just how complex disjunctive theories might have to be to properly account for certain kinds of robust action guidance. Note that his arguments about complexity don't make use of the doubling up problem, which suggests that he is underestimating how complex disjunctive theories will have to be.

14. My thanks to Colin Hickey for helping me to clarify this.

15. We'll discuss an objection to this in Section 4.

one of the *racist* cases Duke does wrong by discriminating, and in the other he does not. Yet in both cases Duke has reasons to apologize, atone, or approximate some ideal, and third parties have reasons to avoid complicity in his racism. On the disjunctive account, it is characteristic of apology, atonement, and avoiding complicity that they are appropriate reactions to wrongdoing. That's part of the attraction of disjunctive theories: they maintain an explanatory connection between obligations and these sorts of robust action guidance. However, given this connection, disjunctive theories have to say that in the *uncontrollable racist* case, Duke and everyone else should treat Duke's behavior as if it is wrong, even though it is not. It's not simply that we should respond to Duke's behavior in ways that are typical of responses to wrongdoing, it's that we should respond in ways characteristically explained by the fact that one has done wrong. This is disunity: there is something disharmonious about telling people to treat acts that aren't wrong as if they were.

Disjunctive theories thus have two unavoidable costs when compared to their closest DOER rivals: complexity and disunity. The doubling up problem is a cost, but an avoidable one; avoiding it, however, either exacerbates the complexity of the theory, or gives up some of its explanatory power. Complexity, disunity, and lack of explanatory power are not enough to show that disjunctive theories are false. However, if we have two otherwise equally good moral theories, we should prefer the more elegant, unified, and explanatory one. Since we cannot appeal to the truth of OIC to adjudicate between DOER and disjunctive theories—that would be begging the question—DOER theories and their closest disjunctive non-DOER competitors are equally good with regard to the prescriptions they issue. DOER theories are more elegant and more unified, and potentially more explanatory, and so should be preferred.

Let's now turn to a type of non-DOER theory that avoids this argument.

3.3. *Elegant Non-DOER Theories*

To be elegant, unified, and also avoid the doubling up problem, a non-DOER theory needs to give the same explanation of robust action guidance in both *ordinary racist* and *uncontrollable racist*. DOER theories appeal to common, morally-relevant features of the cases to explain the obligations in each, and then appeal to obligations to explain robust action guidance. An *elegant non-DOER theory* can use the shared morally-relevant features to explain robust action guidance directly, without obligations playing this mediating role. In both *ordinary* and *uncontrollable racist*, the explanation of why Duke should apologize for his discrimination might be that discrimination is harmful and does not properly respect human dignity. Since our elegant non-DOER theory endorses OIC, these

morally-relevant features can also explain obligations when these obligations are fulfillable. This is as elegant an explanation as that given by the closest DOER theory. And no disunity arises since, on the elegant non-DOER account, robust action guidance has no association with wrongdoing.

On an elegant non-DOER account, obligations never play any role in explaining robust action guidance. On an elegant non-DOER account, when φ would ordinarily be obligatory, but φ cannot be done, it is the collected morally-relevant features of φ that explain any robust action guidance. But these features, rather than the obligation to φ , have to also be the complete explanation of robust action guidance even in cases where there is an obligation to φ which can be fulfilled. If they were not, then elegant non-DOER accounts would lose both their simplicity and their unity, and would run into the doubling up problem. So, according to such a non-DOER theory, when we ask, "Why should so and so apologize?" we can never correctly answer, "Because she did something wrong." When one asks, "Why should I not do such and such?" we can never correctly answer, "Because it will lead you to violate a future obligation." When we ask, "Why shouldn't I benefit from that person's behavior?" we can never correctly answer, "Because his behavior is wrong." It might be true that one's having committed a wrong act explains the need to apologize, but this is just because one's having committed the act explains the need to apologize, and the act happens to be wrong. The wrongfulness of the act plays just as much role in the explanation as the day the act was committed on. Accepting an elegant non-DOER theory thus means giving up the claim that obligations play a central, vital role in guidance. At best obligations would have importance in a very narrow space. The guidance we get from an obligation to φ at time T would just be guidance to φ at T. This guidance is fleeting and of necessity plays only a small role in our lives. Neither the wrongness of past actions, nor the wrongness of others' actions, would ever play a direct role in guidance on such an account. This is a very impoverished view of how obligations guide, at least in the moral realm. Elegant non-DOER accounts allow us to keep OIC, but only by surrendering the heart of the idea that obligations must guide.

Some disjunctive theories give rise to this same problem. The doubling up problem arises because the features that explain robust action guidance when putative obligations cannot be fulfilled also tend to occur when those obligations can be fulfilled. This problem can be addressed by significantly reducing the contribution of obligation violation to robust action guidance, so that there is little difference between *ordinary* and *uncontrollable racist* in the strength of robust action guidance. On this view, obligations play some role in guidance, but only a very minor role. They do very little to explain robust action guidance in any condition, as they always only contribute a small amount of strength to

robust action guidance. This, too, seems to be an impoverished view of how obligations guide.¹⁶

What this shows us is that one of the core motivations that inclined us to accept OIC—the truism that obligations must guide—better supports the denial of OIC.¹⁷ If obligations can guide by explaining robust action guidance, then unfulfillable obligations can guide. Since we can construct DOER and non-DOER theories that give almost exactly the same guidance, we can only decide between them by looking at the virtues of how they explain this guidance. Theories that deny OIC have significant theoretical advantages, in terms of how they explain guidance, over their disjunctive competitors. We can give a non-DOER theory that has all of these advantages, but to do so we have to say that obligations play almost no role in guiding action, and this is really giving up the truism that obligations guide.¹⁸

3.4. Is Robust Action Guidance Really Guidance?

Let's call the principle that obligations must be able to guide action *the guidance principle*. I have argued that DOER theories give a better account of how obligations guide action than non-DOER theories, and so the guidance principle should incline us to deny OIC. My argument only gets off the ground if DOER theories are compatible with the guidance principle; that is, my argument relies on a *broad reading* of the guidance principle, interpreting the principle as saying

16. We can give a non-DOER theory that avoids all the problems I've raised. It would say that in both *racist* cases Duke has violated some fulfillable past obligation, and that this alone is what explains robust action guidance in both cases. To be elegant, the obligation would be the same in both cases. To avoid doubling up, this past obligation would have to be the only obligation in either case. The view would thus say that Duke's discrimination in *ordinary racist* was not wrong. This is too revisionary.

17. I haven't discussed blame and punishment. Some will want to say that Duke should not be blamed or punished in *uncontrollable racist*. Initially, it seems that non-DOER theories can explain this more simply than DOER theories, by saying that he's done no wrong. But this is not so simple, because we then need an explanation for why these forms of robust action guidance are different from other forms that do occur in *uncontrollable racist*. DOER theories can then appeal to the same explanation. This is parallel to my arguments about overguidance in Section 4.

18. Utilitarians will be unconvinced by my arguments. This is because, according to utilitarianism, the values of states are the only sources of reasons to act, and the notion of robust action guidance just can't fit such a theory. I suspect that the lesson to learn from this is that utilitarians should give up on obligations, not that they should endorse OIC. Obligations, on the utilitarian's view, give no more guidance than they do on the elegant non-DOER account. What's more, on any utilitarian account the difference in value between an obligatory act and the next-best non-obligatory act can be vanishingly small. Because utilitarian obligations provide almost no interesting guidance, it's hard to see how such a difference can mark out anything of real moral significance (my ideas here were heavily influenced by Norcross, 2006).

that obligations are able to guide action if they at least provide robust action guidance. My arguments fail if one adopts a *narrow reading* of the guidance principle, one according to which a putative obligation on A to φ at time T is only able to guide action in the right way if it is able to guide A to φ at T. Should we adopt such a narrow reading of the principle? The guidance principle gives us a necessary condition accounts of obligations must meet. If we are to take the narrow reading, rather than a broader one, as giving this necessary condition, we need very good reasons to do so. I will argue in this section that we do not have such reasons.

I would be quite surprised if many of us had intuitions explicitly about the narrow reading of guidance rather than the broad one (although I'm sure that some do), so most of our evidence that directly bears on this is going to come from intuitions about cases, or from an examination of the role obligations and wrongness play in specific instances of action and practical reasoning. This evidence does not favor the narrow reading. It's easy to find cases in which, intuitively, a person is obligated to φ but cannot φ —I've given a few illustrations in Section 2 of common types of cases that elicit such intuitions—and these undermine the narrow reading of the guidance principle. Further, as I will argue in Section 4, even intuitions that seem initially to favor OIC (and thus the narrow reading) are at best equivocal. Turning to our ordinary practices, people often take past or future obligations as giving us reasons to act now (e.g., reasons to avoid or atone), suggesting that we see the function of obligations as providing a broad sort of guidance. This isn't decisive evidence in favor of a broad reading of the guidance principle. If we had some other strong reason to favor the narrow reading, we could likely interpret all of this to be consistent with that narrow reading. But our intuitions and ordinary practices do not by themselves provide good evidence in favor of the narrow reading of the guidance principle.

The other way to justify endorsing the narrow reading of the guidance principle as a necessary condition for accounts of obligation would be to show that our account of obligations gains something very important only if the narrow reading is true. Important things do turn on the truth of the guidance principle, but we seem to gain nothing of significance if it is the narrow reading, rather than some broader version, that is true. We can see this by considering some of the main motivations given by those who endorse the guidance principle. The guidance principle is motivated sometimes by the claim that obligations that cannot guide are "pointless" (e.g., Copp 2003). This favors a broad reading of the guidance principle, not a narrow one. Obligations that cannot guide would only be pointless if they could not guide in a broad sense—if they could not make a difference to *anyone's* actions. This is because (as we've seen) there is a point to obligations which cannot be fulfilled yet still provide robust action

guidance. Another motivation for the guidance principle is that obligations are a kind of reason, and it is characteristic of reasons that we can use them in reasoning about what to do (e.g., Searle 2001; Vranas 2007); on some views this is what distinguishes the reasons for an action from the goodness of that action (Williams 1995; Kolodny 2005).¹⁹ This motivation is consistent with both a broad and narrow reading of the guidance principle. If we accept a DOER theory, unfulfillable obligations can play a role in reasoning about what to do: when one reasons about what to do now, one must take into consideration one's past or future unfulfillable obligations and the unfulfillable obligations of others.²⁰ Further, DOER theories can distinguish reasons or obligations for an action from the goodness of that action, as long as the obligation to φ generates robust action guidance and the goodness of φ -ing does not directly do so (although goodness may explain some obligations, and thus indirectly explain robust action guidance). There thus seems to be no benefit to adopting a narrow reading of the guidance principle rather than a broader one.

What's more, the arguments I have made in this paper suggest that we lose something by adopting a narrow reading of the guidance principle. The narrow reading entails OIC, and means that the correct theory of obligations is either an elegant non-DOER theory or a disjunctive non-DOER theory. As I showed in the previous sub-section, elegant non-DOER theories give wrongness and obligations an impoverished role in action and practical reasoning. On these theories, the only choice that A's obligation to φ at T is relevant to is A's choice at T. All deliberation at all other times and about all other decisions can simply ignore this obligation. This view, which says that every particular obligation is only of fleeting importance, seems to me to be giving up something vital. Those moved towards disjunctive non-DOER theories also recognize this as a loss; that's why they want to allow obligations to play a broader role in action and practical reasoning by saying that they sometimes explain robust action guidance. But then why insist on the narrow reading of the guidance principle? The evidence doesn't give much support for it, and it doesn't seem to give our theory anything of value.

19. Kolodny and Williams are discussing internal/external reasons, not OIC, but their arguments clearly connect to the guidance principle as well.

20. One might insist that this motivation for the guidance principle is not consistent with a broad reading, because (one would claim) what characterizes reasons to φ is that they can be used in reasoning about *whether to* φ . But, since this is not obviously true, and (as I will argue) intuitions don't really support it at all, one would have to show that something important is gained by insisting on this. If we just want to make sure reasons have a point, or to distinguish reasons from good-making features, we can do this without adopting this narrow view.

4. Intuitions and Overguidance

Let's turn now from arguments based on general principles to arguments based on intuitions about cases. With the assistance of the notion of robust action guidance, we will see that the intuitions which seem most clearly to support OIC do no such thing.

As above, it will be helpful to use two parallel examples to frame our discussion. In the previous section, we considered examples in which an agent could and could not control certain behaviors, and discussed how to account for the fact that similar robust action guidance could arise in both cases. In this section we'll use parallel cases in which an agent can and cannot control certain behaviors, but where that *does* seem to make a difference to robust action guidance. Here are the two examples I'll focus on:

avoidable baby: A baby is drowning in a shallow pond. Daisy is the only person around, and knows the baby is drowning. Nothing prevents her from saving the baby, but she does not save it.

unavoidable baby: This is exactly like *avoidable baby*, except that Daisy is trapped under a fallen tree, through no fault of her own. She fails to save the drowning baby.

Intuitively, Daisy violates an obligation in *avoidable*, but not *unavoidable baby*. Equally importantly, there is an intuitive difference in the robust action guidance Daisy receives in each case. In *avoidable baby*, Daisy should apologize and atone for what she has failed to do. However, the same is not true in *unavoidable baby*: Daisy simply does not owe an apology or restitution for her failure to save the baby.²¹

Intuitions about these cases seem to pose a problem for theories which deny OIC (Howard-Snyder 2006; Vranas 2007). A DOER theory needs some way of explaining why Daisy does wrong in *avoidable baby* but not in *unavoidable baby*. It can't straightforwardly appeal to her inability, since DOER theories deny OIC. Nor can it bite the bullet and say she ought to save the baby in both cases, because this would generate inappropriate robust action guidance. Let's coin a term to help us talk about this. The term is *overguidance*. A theory overguides when it assigns obligations or robust action guidance in cases where they are clearly inappropriate. DOER theories overguide if they say that Daisy was ob-

21. To be clear, she may have some reasons to apologize in *unavoidable baby*. Her apology or acts of restitution might help the family through its grieving process. But her reasons to help the family wouldn't constitute robust action guidance, since they would not really arise in response to Daisy's failure. If someone completely unrelated to the drowning could help the family in the same way by apologizing, they would have the same reasons.

ligated to save the baby, or that she should apologize or atone after she fails, in *unavoidable baby*. Let's call the intuitions that such overguidance is inappropriate *overguidance intuitions*. Almost any appeal to intuitions about specific cases to support OIC is going to appeal to overguidance intuitions. The cases that separate close competing DOER and non-DOER views are not those where an agent *can* fulfill some putative obligation, since the disagreement on OIC won't be relevant here. So the cases have to involve agents who can't fulfill putative obligations, and only intuitions that the agent lacks the putative obligation or should not be robustly guided will support OIC. If it turns out that overguidance intuitions don't favor non-DOER theories over DOER theories, this will seriously weaken the case for OIC.

Arguments for OIC sometimes appeal to overguidance intuitions about more bizarre obligations than those in *unavoidable baby* (e.g., Howard-Snyder 2006; Vranas 2007). For example, it would be great, morally speaking, if I were to fly around the world saving babies like Superman. But this is clearly not obligatory, and I do not owe anyone an apology if I fail to do this. DOER theories must be able to agree. In what follows, I am going to focus just on the *avoidable baby* and *unavoidable baby* cases, because if DOER theories can explain why Daisy lacks the relatively ordinary obligation to save the baby when she is trapped under the tree, they can also explain why we lack the obligations to be superheroes.

Surprisingly, overguidance intuitions are also a problem for theories that accept OIC. A non-DOER theory can say that Daisy does nothing wrong in *unavoidable baby*, but this by itself doesn't avoid overguidance. This is because, if a non-DOER theory had to say that Daisy should apologize for her failure to save the baby in that case, this would also conflict with our overguidance intuitions. As we've seen, non-DOER theories must say that robust action guidance can and does arise in response to actions that were not wrong—even in response to actions that were not wrong because they were outside the agent's control. This is the lesson of the *racist* cases. Put another way, the *racist* cases show us that the inability to ϕ does not by itself mean that robust action guidance does not arise from the failure to ϕ . And so non-DOER theories can't straightforwardly point to Daisy's inability to save the baby in *unavoidable baby* to explain why she need not apologize or atone.

Both DOER and non-DOER theories need a response to overguidance intuitions. DOER theories need some way of explaining why Daisy was not obligated to save the baby, and thus has no reasons to apologize for not doing so, in *unavoidable baby*. Non-DOER theories need some way of explaining why Daisy has no reasons to apologize or atone in *unavoidable baby*. There are two ways of trying to turn overguidance intuitions into an argument for OIC. First, one might claim that only non-DOER theories can fit these intuitions—that every DOER theory says that Daisy is obligated to save the baby in *unavoidable baby*.

Alternately, one might claim that any solutions to the problem a DOER theory can employ will undermine the arguments I made in the previous sections.²² The idea behind this alternate route is that we must add something to DOER theories to explain the absence of robust action guidance in *unavoidable baby*, and this makes DOER theories less elegant than I have claimed they are. This would significantly weaken the argument in favor of DOER theories over disjunctive theories. And it might give an argument in favor of elegant non-DOER theories, if it turns out that they are more elegant than their DOER rivals.

These arguments based on overguidance intuitions might work against some DOER theories, but not all. There are DOER theories that can accommodate overguidance intuitions. What's more, for at least some DOER theories, the explanation of the difference between the two *baby* cases falls out of their already simple account of obligations. These theories require no additional complexity to deal with overguidance, and are thus no less elegant than their close non-DOER competitors. Finally, even if we must add some complexity to a DOER theory to deal with the overguidance problem, it doesn't look like any elegant non-DOER theories can ever give a simpler solution than their closest DOER competitors. Disjunctive theories may be in even worse shape. I'll argue for these points first via illustration, and then make some more general arguments.

Consider a DOER version of a Kant-inspired deontology. This view says one is obligated to will only in accordance with the categorical imperative; that is, it is wrong to will to do an act that is not in accordance with a universalizable maxim (or that doesn't treat others as ends). Let's assume that we are not fully in control of our will, as I suspect empirical psychology tells us. If so, some agents will be unable to always will appropriately, and OIC will be false. For example, Duke's actions are wrong in both the *racist* cases because he wills his discrimination in both cases, even if in *uncontrollable racist* he is unable to will to not discriminate. This view gets its solution to the overguidance problem for free: it simply follows from the account of where obligations come from, without any special complexity being needed to deal with it. Daisy violates no obligation in the *unavoidable baby* case, since there's no plausible construal of the categorical imperative according to which Daisy wills improperly while trapped under a tree. So no robust action guidance arises in this version of the case. The elegant non-DOER version of this view is no simpler. Like all elegant non-DOER theories, it will appeal to the morally-relevant features of acts and omissions to explain robust action guidance. So it will say that robust action guidance is generated when an agent fails to will in accord with the categorical imperative. The only difference between this and the DOER version is that on the non-DOER version an agent

22. My thanks to Chris Heathwood and Tom Dougherty for helping me properly understand this aspect of the overguidance objection.

will only be *obligated* so to will when they are able to do so. This does not make this elegant non-DOER view any more elegant than the DOER version. Finally, let's consider the disjunctive non-DOER version of the view. According to such a view, obligations always generate robust action guidance. However, robust action guidance can also be explained in some other way. The obvious option is that it also arises from failures of will. This generates a doubling up problem: an agent who can properly will, but does not, both does wrong and fails to properly will. Since both can generate robust action guidance, it seems that robust action guidance to approximate the ideal, atone, or avoid complicity should be weaker in *uncontrollable racist* than in *ordinary racist*, which is implausible. The disjunctive Kantian view needs some feature that deals with this problem, and that should reduce its elegance.

We can see much the same results with DOER versions of at least some other familiar theories.²³ But it will be more helpful to us to consider a DOER view that does bring in additional complexity to deal with the overguidance problem. This view is suggested by Peter Graham (2011). The view distinguishes between negative and positive duties. Negative duties are duties to refrain from certain acts, such as discrimination. Positive duties are duties to do certain things, such as save babies. OIC is false on this view: the view allows that there can be unfulfillable obligations. However, the view says that only negative duties can be unfulfillable. This is because all (and only) positive duties have "if possible" built into their content. That is, the duty to save drowning babies is really the duty to save drowning babies if one can. On the other hand, the duty to not discriminate is a negative duty and so contains no "if possible" proviso. So it would be wrong to discriminate even if one could not avoid discriminating. If this view were a DOER theory, it could say that wrong is done, and robust action guidance generated, in *uncontrollable racist* but not in *unavoidable baby*. This view includes machinery to specifically deal with the overguidance problem, and this adds some complexity to it. However, the elegant non-DOER version of Graham's view seems to be at least as complex as the DOER version. The elegant non-DOER version of Graham's view would say that there are morally-relevant features associated with any negative duty to φ that are still present when φ cannot be done, and it is these that explain robust action guidance. The view must also say that the morally-relevant features associated with positive duties to ψ , and which explain robust action guidance, do not persist when ψ cannot be done. This looks no simpler than the DOER version, since both need some way of differentiating

23. For example, a DOER quasi-virtue ethics might say that one ought to do whatever the ideal agent would do if she found herself in one's circumstances; the ideal version of Daisy still wouldn't save the baby in *unavoidable baby*, but the ideal version of Duke would refrain from discriminating. The solution to the overguidance problem just falls out of the theory, and the non-DOER versions can be no simpler.

between cases that involve what would ordinarily be positive and negative duties. Compare these to the disjunctive non-DOER version of Graham's theory, which says that robust action guidance comes either from obligations or from some other source. Just like the DOER version, the disjunctive theory must explain why robust action guidance gets generated in *uncontrollable racist* but not *unavoidable baby*, and if it appeals to the difference between positive and negative duties, it seems to have no advantage over the DOER version. Further, the theory must contain two sources of robust action guidance—obligations, and whatever explains guidance when there is no fulfillable obligation. And the theory must explain why robust action guidance does not double up in *ordinary racist*. When it comes to elegance, Graham's view seems likely to have the advantage over close disjunctive alternatives.

What do we learn from these examples? Intuitions about overguidance can be raised as objections to theories which accept OIC as well as against those that deny it. At least some theories that deny OIC seem to be able to fit these intuitions. One might try to adapt the overguidance worry into an argument that DOER theories are less elegant than their non-DOER rivals. But this will not clearly work. Any solution to the overguidance problem an elegant non-DOER theory can give should also be available to the closest DOER theory. This is because the only necessary difference between DOER and elegant non-DOER theories is the acceptance of OIC, and OIC can't by itself explain when robust action guidance does or does not arise (if it did, we'd get the wrong guidance in *uncontrollable racist*). So, an elegant non-DOER theory will have to appeal to something else about the morally-relevant features of the *baby* cases to explain the difference in robust action guidance. Whatever else it appeals to should be something that the closest DOER theory can appeal to as well.²⁴ Therefore, the closest DOER theory should be able to explain the difference in obligations by the same sort of appeal to Daisy's ability. Overguidance intuitions do not give us grounds to decide between DOER and elegant non-DOER theories. What's more, DOER theories potentially do better than their disjunctive non-DOER rivals at fitting overguidance intuitions. Some DOER theories have an account of obligations which straightforwardly fits intuitions, and adapting this account into a disjunctive theory generates the doubling up problem, and complexity is needed to deal with this. It is likely we can find many such DOER theories, for example by giving more sophisticated versions of the simple theories just discussed. And at least some DOER theories which must add complexity to deal

24. It may be that facts about Daisy's abilities play a role in the explanation of the difference between the cases. But we can appeal to facts about ability without appealing to OIC. The DOER virtue ethics mentioned in Footnote 23 explains why Daisy does not get robust action guidance in *unavoidable baby* by pointing out that even the ideal version of Daisy can't get out from under a tree.

with overguidance intuitions still seem to come out less complex than competing disjunctive theories.

Let's summarize what we have seen. Intuitions about cases are often used to argue that OIC is true. However, these intuitions do nothing to decide between certain rival DOER and elegant non-DOER theories. What's more, intuitions about cases favor these DOER theories over rival disjunctive non-DOER theories, because the overguidance problem seems to exacerbate the theoretical vices of disjunctive non-DOER theories. We thus cannot give a non-DOER theory that fits the notion that obligations must guide, fits our intuitions about cases, and is theoretically virtuous. Elegant theories do the last two well, and disjunctive theories do the first two well. At least some DOER theories do all three. This is not conclusive. It may turn out that, once we do more to figure out what the correct moral theory is, some non-DOER version of that theory will have a better solution to the overguidance problem than the DOER version. But none of the supposed best evidence for OIC actually gives us reason to expect the true moral theory to accept OIC.²⁵

5. What about Reasons?

Many philosophers think that oughts and obligations are generated by reasons. If they are right, then if there are unfulfillable obligations, agents sometimes have reasons that cannot be acted on. Some find this implausible (e.g., Vranas 2007; Streumer 2007; see Heuer 2010 for responses, however). I'll briefly say how the arguments given in this paper can shed light on this issue.

Reasons which can't be acted on might be problematic because reasons should be able to guide agents. I've argued that unfulfillable obligations can guide, and so the reasons which generate these obligations guide as well. But the worry about guidance is not so easily responded to. To see why, consider first a different worry: if we can have reasons that we can't act on, then we should have reasons to do things like fly around the world solving problems like Superman. To some, this seems "absurd" (this is Streumer's, 2007, word). It seems to me, though, that as long as these reasons don't generate absurd obligations, then the

25. David Copp (2003) has argued for OIC by claiming that it would be unfair if someone were obligated to do acts they could not do, and that morality cannot be unfair. The arguments I have given throughout this paper show that worries about unfairness do not push us to accept OIC. Non-DOER theories must still require agents to apologize or make restitution for acts they could not help doing. As we see from the discussion of overguidance, DOER and non-DOER theories need not differ in when they require this. Requiring apologies and restitution for acts which are not wrong, but which one could not help but do, is no fairer than saying these acts are wrong and also requiring apologies or restitution.

existence of these reasons is not so absurd (see Crisp, 2006, for agreement; some of Streumer can be read as sympathetic to this response). My arguments about overguidance show that, even if we have reasons to act like Superman, these need not generate absurd obligations. But this brings us back to the guidance problem: if we have reasons to act like Superman, then these reasons provide no guidance at all.

DOER theories can say one of two things. They can accept that we have reasons to act like Superman, and thus that reasons do not always guide, or deny that we have these reasons. While I think the former is not so bad (as long as obligations always guide), let's focus on the latter. How could one deny that we have reasons to do the extremely good things Superman would do? One motivation for the claim that reasons must be able to guide is the view that reasons must be the sort of thing that we appeal to when reasoning about what to do (Searle 2001; Streumer 2007). On a DOER theory, it would make sense to reason about some impossible act φ —that is, to consider reasons to φ even though one cannot φ —as long as φ might turn out to be obligatory. This is because, if φ is obligatory, then the failure to φ generates robust action guidance, which is important to know when deliberating about what other acts to do (e.g., whether to apologize for not φ -ing). However, when φ is impossible and there is no chance that φ is obligatory, it makes no sense to engage in reasoning about φ . This might give a DOER theory grounds for claiming that there can be reasons to φ when φ is impossible, as long as φ is potentially obligatory, but when φ is neither potentially obligatory nor possible there can be no reasons to φ . This would allow one to deny that we have reasons to fly around the world solving problems like Superman while still thinking that we have reasons to do some things we cannot do. Thus, reasons can always guide even if we can have reasons to do things we cannot do.

6. Conclusion

We started with the truism that oughts and obligations must guide. Some might take this to mean that an obligation to φ does nothing more than tell us to φ . This is the elegant non-DOER view. However, the truism is better understood as saying that obligations are at the center of our normative lives, that the wrongness of acts or omissions must play an important role in explaining how we and others should respond to these acts or omissions. Elegant non-DOER views cannot fit this idea. Only disjunctive non-DOER views and views that reject OIC can. Theories that reject OIC give theoretically superior accounts of robust action guidance than disjunctive non-DOER theories. Theories that reject OIC also give superior responses to the overguidance worry—the disadvantages of disjunctive

non-DOER theories seem worsened by the need to fit overguidance intuitions. If oughts and obligations guide, then we should think that some obligations cannot be fulfilled.

Acknowledgements

I greatly appreciate the help I got with this paper from Dom Bailey, Eric Chwang, Chris Heathwood, Colin Hickey, Eric Hochstein, Kathryn Lindeman, Daniel Nolan, Nick Southwood, Justin Snedegar, Julia Staffel, and Eric Wiland. I would also like to thank many helpful conference attendees and journal referees whose names I do not know. Finally, I could not have written this paper without having discussed these ideas with my students over the years, and I am grateful for that opportunity.

References

- Baker, Alan (2013). Simplicity. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Fall 2013 Edition)*. Retrieved from <http://plato.stanford.edu/archives/fall2013/entries/simplicity/>
- Brown, James (1977). Moral Theory and the Ought-Can Principle. *Mind*, 86(342), 206–223. <http://dx.doi.org/10.1093/mind/LXXXVI.342.206>
- Copp, David (2003). 'Ought' Implies 'Can', Blameworthiness, and the Principle of Alternate Possibilities. In David Widerker and Michael McKenna (Eds.), *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (265–299). Ashgate.
- Crisp, Roger (2006). *Reasons and the Good*. Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199290338.001.0001>
- Gendler, Tamar S. (2008). Alief in Action (and Reaction). *Mind & Language*, 23(5), 552–585. <http://dx.doi.org/10.1111/j.1468-0017.2008.00352.x>
- Gendler, Tamar S. (2011). On the Epistemic Costs of Implicit Bias. *Philosophical Studies*, 156(1), 33–63. <http://dx.doi.org/10.1007/s11098-011-9801-7>
- Goble, Lou (2009). Normative Conflicts and the Logic of 'Ought'. *Noûs*, 43(3), 450–489. <http://dx.doi.org/10.1111/j.1468-0068.2009.00714.x>
- Graham, Peter A. (2011). 'Ought' and Ability. *Philosophical Review*, 120(3), 337–382. <http://dx.doi.org/10.1215/00318108-1263674>
- Hare, Richard M. (1963). *Freedom and Reason*. Oxford University Press.
- Heuer, Ulrike (2010). Reasons and Impossibility. *Philosophical Studies*, 147(2), 235–246. <http://dx.doi.org/10.1007/s11098-008-9285-2>
- Howard-Snyder, Frances (2006). "Cannot" Implies "Not Ought." *Philosophical Studies*, 130(2), 233–246. <http://dx.doi.org/10.1007/s11098-004-4511-z>
- Kolodny, Nico (2005). Why Be Rational? *Mind*, 114(455), 509–563. <http://dx.doi.org/10.1093/mind/fzi509>

- Lowry, Rosemary (2012). Reasons for Action and Psychological Capacities. *Ethical Theory and Moral Practice*, 15(4), 521–531. <http://dx.doi.org/10.1007/s10677-011-9307-6>
- Marcus, Ruth B. (1980). Moral Dilemmas and Consistency. *The Journal of Philosophy*, 77(3), 121–136. <http://dx.doi.org/10.2307/2025665>
- Norcross, Alastair (2006). Reasons without Demands: Rethinking Rightness. In James Dreier (Ed.), *Contemporary Debates in Moral Theory* (38–53). Blackwell Publishing.
- Ryan, Sharon (2003). Doxastic Compatibilism and the Ethics of Belief. *Philosophical Studies*, 114(1), 47–79. <http://dx.doi.org/10.1023/A:1024409201289>
- Searle, John (2001). *Rationality in Action*. MIT Press.
- Singer, Peter (1972). Famine, Affluence and Morality. *Philosophy and Public Affairs*, 1, 229–43.
- Sinnott-Armstrong, Walter (1984). “Ought” Conversationally Implies “Can.” *The Philosophical Review*, 93(2), 249–261. <http://dx.doi.org/10.2307/2184585>
- Staffel, Julia (2015). Measuring the Overall Incoherence of Credence Functions. *Synthese*, 192(5), 1467–1493. <http://dx.doi.org/10.1007/s11229-014-0640-x>
- Streumer, Bart (2007). Reasons and Impossibility. *Philosophical Studies*, 136(3), 351–384. <http://dx.doi.org/10.1007/s11098-005-4282-1>
- Streumer, Bart (in press). Reasons and ability. In Daniel Star (Ed.), *Oxford Handbook of Reasons and Normativity*.
- Vranas, Peter B. (2007). I Ought, Therefore I Can. *Philosophical Studies*, 136(2), 167–216. <http://dx.doi.org/10.1007/s11098-007-9071-6>
- Williams, Bernard (1995). Internal Reasons and the Obscurity of Blame. In Bernard Williams, (Ed.), *Making Sense of Humanity: And Other Philosophical Papers 1982-1993*. (Reprinted from *Reason and Moral Judgment*, Logos, Vol. 10, by James Prior, Ed., 1989). <http://dx.doi.org/10.1017/cbo9780511621246.004>
- Zimmerman, Michael J. (1996). *The Concept of Moral Obligation*. Cambridge University Press. <http://dx.doi.org/10.1017/cbo9780511624681>