

## Free Will and the Asymmetrical Justifiability of Holding Morally Responsible

Ben Vilhauer, William Paterson University of New Jersey

### Introduction

Among Susan Wolf's contributions to the free will debate is to highlight the intuition that we must meet a higher standard of justification to blame than to praise.<sup>1</sup> Wolf points out this intuition in the course of arguing that the control condition of moral responsibility is different in the contexts of praise and blame. (On her view, people can be justifiably blamed only if they could have done otherwise, but can be justifiably praised even if they could not have done otherwise.) But apart from Wolf's work and responses to it, there seems to have been little discussion of this intuition. This is surprising, because it seems to have interesting implications for free will theory even if one does not follow Wolf in holding that the control condition of moral responsibility is different in the contexts of praise and blame. (In this paper, "free will" means whatever satisfies the control condition of moral responsibility, and "acting freely" means acting with free will.<sup>2</sup>) That is, it seems to imply that even if we accept a theory on which the control condition is the same in the contexts of praise and blame, we must have better reasons to believe that the theory is true when we appeal to it to justify blame than when we appeal to it to justify praise. For example, even if we accept a Frankfurt-style compatibilism<sup>3</sup> about the control condition in the contexts of both praise and blame, this intuition seems to imply that we must have better reasons to believe that Frankfurt-style compatibilism is true to appeal to it to justify blame than to justify praise. The same would appear to hold if we instead accepted a dual-control libertarian account of the control condition in the contexts of both praise and blame. My goal in this paper is to develop this point, and then to draw some tentative conclusions about when (if ever) we can justify holding people morally responsible. The idea is to see how far this

project can be carried without making independent commitments on the other issues that typically occupy the attention of free will theorists (such as whether compatibilism is true and whether free will requires alternative possibilities or agent causation).

This paper has six sections. In the first, I discuss the intuition about the asymmetric justifiability of praise and blame in more detail. In the second, I explain why the asymmetric justifiability of praise and blame implies an asymmetry in the justificatory standards that must be met by claims about free will in the contexts of praise and blame. I then ask: how much higher is the standard for blame? In sections three and four I offer a partial answer: seriously harmful blame can only be justified if it cannot be reasonably doubted that the target of blame had free will, but praise which can be given at no significant cost to anyone can be justified as long as it can be reasonably doubted that the candidate for praise did not have free will. In the fifth section, I argue that anyone who thinks the debate about whether we have free will is philosophically valuable must accept that reasonable doubt is possible in both these cases. In the sixth, I comment further on the difference between Wolf's view and mine.

### 1. The Asymmetry Intuition

In this paper, "blame" will be used to refer to actions used to express the negative reactive attitudes, and "praise" will be similarly used to refer to actions used to express the positive reactive attitudes. Understood in this way, blame and praise involve not just *believing* that someone is morally responsible, but acting in such a way as to *hold* him morally responsible.

The intuition that praise and blame are asymmetrically justifiable can be explained as part of a more general intuition that harms and benefits are asymmetrically justifiable. Justice demands that arguments for harming people be to be held to a higher standard than arguments for

refraining from harming them or benefiting them. All philosophers must acknowledge that this asymmetry exists, though disagreement is to be expected when it comes to giving a detailed explanation of why it exists. All will probably agree that, one way or another, the primary purpose of morality has to do with getting people to benefit and refrain from harming each other. If this is right, then arguments for harming people run against the primary purpose of morality (in at least a *prima facie* way), while arguments for benefiting people run with it. In my own view, this asymmetry exists because of personhood-based desert claims which can be made by or on behalf of all people. In other words, it exists because of how people deserve to be treated just by virtue of being people. People deserve not to be intentionally harmed unless there is a very strong justification for that harm. People also deserve to be benefited when someone can do so at no significant cost to anyone. When someone is considering providing such a benefit, the burden of justification is much lighter: that is, if all parties to an interaction agree that something is a benefit which can be provided at no significant cost to anyone, no further justification is typically expected.

The strength of the desert claims people can make may be different in the cases of harm and benefit. We fail to live up to the demands of morality whenever we intentionally harm someone without a very strong justification. But it may sometimes be possible to refrain from benefiting people without failing to live up to the demands of morality, even in cases where benefits could be provided at no significant cost to anyone. But no matter how one sees the details on this point, it is clear that there is a moral presumption in favor of refraining when it comes to harm, and in favor of acting when it comes to benefit. Put differently, if everyone agrees that some action under consideration would be a harm, the burden of proof is on whomever claims that this action would be legitimate. But if everyone agrees that some action

under consideration would be a benefit that could be provided at no significant cost to anyone, the burden of proof would be on anyone who claims that it would be wrong to act in this way.

When praise and blame are successful, that is, when they have their intended effects, they are kinds of benefit and harm. That is, even in cases where praise and blame do not include the obvious benefit and harm of reward and punishment, they are intended to cause valuable or harmful emotional responses in the people who are their recipients. So the asymmetry in the justificatory standards for harm and benefit seems to imply an asymmetry in the justificatory standards for praise and blame.

Praise and blame are special kinds of benefit and harm, since they can only be legitimate if the people who are their recipients deserve them based on how they have acted. We owe it to people to treat them in accordance with the justificatory asymmetry just because they are people. But if people ever really deserve praise or blame, then they deserve it not just because they are people, but because of how they have acted. But someone can only deserve something based on his actions if he is morally responsible for those actions.

My own view is that we need to distinguish between personhood-based desert and action-based desert to explain this difference. (The idea is that all acts of harm and benefit must conform to personhood-based desert, but praise and blame are special in that they must also conform to another kind of desert, action-based desert.<sup>4</sup>) But for purposes of this paper, the details of how one explains the difference are probably less important than seeing the difference that needs explaining.

Some might wonder whether the claims about moral responsibility involved in justifications of praise and blame make praise and blame so different from other kinds of harm and benefit that the justificatory asymmetry relevant for other kinds of harm and benefit has no

bearing on the justifications for praise and blame. But I cannot see why this should be so. It is a straightforward matter to include claims about moral responsibility within the scope of the justificatory asymmetry. We can hold justifications for blame to a higher standard than justifications for praise by holding all the claims that play roles in justifications of blame to a higher standard than the claims that play roles in justifications of praise. Since claims about moral responsibility play roles in justifications of both praise and blame, those claims must be held to a higher standard when they appear in justifications for blame than when they appear in justifications for praise.

How much higher is the justificatory standard for claims about moral responsibility in justifications for blame than in justifications for praise? If the praise/blame asymmetry exists because of the harm/benefit asymmetry, then it seems natural to suppose that the standard for praise is a slope that rises as the cost of that praise increases, and that the standard for blame is a slope that rises as the harmfulness of blame increases. But it is beyond the reach of this paper to survey this whole continuum. Here the argument will focus on the standards for two kinds of cases at the extremes of the continuum: seriously harmful blame, and praise that can be given at no significant cost to anyone.

It is probably clear enough how blame can be seriously harmful. Blame can include retributive bodily violence. Blame can also include the retributive infliction of great emotional pain as punishment for bad actions, and this emotional violence can sometimes be even more harmful than bodily violence.<sup>5</sup> Other kinds of seriously harmful blame include retributive punishment by execution, and by imprisonment under dreadful conditions that make rehabilitation all but impossible (like those that prevail in contemporary prisons).<sup>6</sup>

The idea of praise which can be given at no significant cost to anyone is probably less familiar. What I have in mind is praise that benefits the recipient and doesn't do any significant harm to anyone. For example, it doesn't make the person giving the praise miss a significant opportunity to do something else worthwhile. Also, there are no third parties who are harmed in any significant way by the praise. That is, if we select one person from a group of people to praise her for doing something good, this can cause the other people in the group to feel sorrowful about not having done well enough to merit recognition. The valuable experience of the person singled out comes at the expense of the painful experiences of the others. But there are surely possible instances where costs like these could be avoided. There are probably some actual instances too. We can sometimes single out individuals privately so that others' feelings are not hurt. On other occasions, we can be egalitarian with praise. Suppose everyone in the group has tried to do good things. We might praise them all for trying.<sup>7</sup> It may well be that everyone who can be hurt by being excluded from praise sometimes tries to do good things. (That is, it may well be that everyone sometimes tries to do good things except sociopaths who do not care whether or not they are praised.) If this is right, then we might maximize the amount of costless praise we give by distributing praise in an egalitarian way among everyone we meet who sometimes tries to do good things.

In what follows, for the sake of brevity, seriously harmful blame will be referred to as "serious blame", and praise that can be given to someone at no significant cost to anyone will be referred to as "costless praise".

## 2. Asymmetry and Free Will

As said earlier, praise and blame are benefits and harms which can only be legitimate if the people who are their recipients are morally responsible for the actions at issue. To be morally responsible, one must satisfy the conditions of moral responsibility. There is broad agreement that there are at least two conditions of moral responsibility, an epistemic condition and a control condition. The focus in this paper is on the control condition. But for present purposes, it is important to draw further distinctions between three kinds of conditions which are important to our thinking about control and moral responsibility:

- (i) the control condition of moral responsibility
- (ii) the condition of a justified *belief* that someone satisfies the control condition
- (iii) the condition of being justified in *acting* in a way that treats someone as satisfying the control condition of moral responsibility.

Condition (i) is satisfied by anyone who acted freely. Condition (ii) is satisfied not by the fact that someone acted freely, but rather by a justified belief that he acted freely. Condition (iii) is also satisfied by a belief, but I will argue that it is different in the contexts of serious blame and costless praise. More specifically, I will argue that (iii) is satisfied for serious blame only if it cannot be reasonably doubted that the target of the blame acted freely, and that it is satisfied for costless praise so long as it can be reasonably doubted that the candidate for the praise did not act freely.

In this paper, it will only be argued that there is an asymmetry at level (iii), not at level (i) or (ii). An asymmetry at level (i) seems to imply an asymmetry at levels (ii) and (iii), and an asymmetry at level (ii) seems to imply an asymmetry at level (iii). But an asymmetry at level (iii) does not imply an asymmetry at level (i), and probably does not imply asymmetry at level (ii).

As mentioned at the outset, Susan Wolf holds that there is an asymmetry at level (i). On her view, alternative possibilities are necessary to satisfy the control condition of blameworthiness, but not praiseworthiness. This level (i) asymmetry implies a level (ii) asymmetry. That is, if alternative possibilities are necessary to satisfy the control condition for blameworthiness but not praiseworthiness, then a justified belief that someone had alternative possibilities is necessary for a justified belief that he satisfies the control condition of blameworthiness but not praiseworthiness. This in turn seems to imply an asymmetry at level (iii). If a justified belief that someone had alternative possibilities is required for a justified belief that he satisfies the control condition of blameworthiness but not praiseworthiness, then we presumably need some reason to believe that he had alternative possibilities to treat him as satisfying the control condition of blameworthiness, but not to treat him as satisfying the control condition of praiseworthiness.

Wolf's view that there is a level (i) asymmetry has been discussed with great interest in the literature, but it remains a minority view. I do not claim that this view is false, only that there can be a level (iii) asymmetry without a level (i) asymmetry. Suppose that to justifiably believe that something is the case, we must have reason to think that there is a greater than 50% chance that it is the case. (This is controversial, of course, but this argument does not turn on the precise number we choose here—it could easily range between 50% and 75%, for example). Next suppose that we justifiably believe that there is no level (i) asymmetry, that is, that the very same kind of control of one's actions satisfies the control condition of moral responsibility in the contexts of praise and blame. Also suppose that we justifiably believe that the kind of control which satisfies the control condition in both contexts is control of kind C. In these circumstances, we should probably hold that there is no level (ii) asymmetry. That is, we should

probably hold that reason to think there is a greater than 50% probability that someone possessed C justify the belief that he possessed C in the contexts of both praise and blame.

However, we might think that the symmetry disappears when we arrive at level (iii), that is, when it comes to actually *holding* people morally responsible, especially when we consider the extremes of costless praise and serious blame. (We might think that beliefs are morally neutral when they are not expressed in action, so the justificatory asymmetry is only relevant when actions of holding morally responsible are at issue.<sup>8</sup>) If condition (iii) is satisfied in the context of serious blame only if there is reason to think that the probability that the target acted freely which is substantially higher than 50%, then it can fail to be satisfied in cases where the condition of believing that he deserves serious blame is satisfied. Suppose (as will be argued below in section 3) that condition (iii) is satisfied for serious blame only if it cannot be reasonably doubted that the target acted freely. We might have justified beliefs that C satisfies the control condition of moral responsibility, and that the target possessed C, but if either of these claims could be reasonably doubted, then the claim that the target acted freely could be reasonably doubted, and condition (iii) would not be satisfied.

Now consider costless praise. Suppose (as will be argued below in section 4) that condition (iii) is satisfied for costless praise as long as it can be reasonably doubted that the candidate did not act freely. If this is right, then it can be satisfied in cases where the condition of believing that he deserves costless praise is not satisfied. We might not have justified beliefs that C satisfies the control condition of moral responsibility, or that the candidate possessed C, but if the negation of both of these claims can be reasonably doubted, then the claim that the candidate did not act freely can be reasonably doubted, and condition (iii) would be satisfied for costless praise.

### 3. Serious Blame and Reasonable Doubt

In this section, it will be argued that condition (iii) is satisfied for serious blame only if it cannot be reasonably doubted that the target acted freely.<sup>9</sup> Call this the *serious blame principle*. The serious blame principle implies that there is an obligation to refrain from serious blame whenever it can be reasonably doubted that the target of the blame had free will with respect to the action at issue. In other words, according to the serious blame principle, if it can be reasonably doubted that the target had free will, condition (iii) is not satisfied for serious blame, so serious blame is not justified. If serious blame is not justified, then there is an obligation to refrain from it.

The serious blame principle should be accepted because of its close kinship to another "reasonable doubt" principle which is widely recognized to be a requirement of justice, that is, the requirement in criminal legal proceedings that the accused can only be convicted of a crime if it is proven beyond reasonable doubt that he acted criminally. The grounds for accepting the reasonable doubt criminal conviction standard are the same as the grounds for accepting the serious blame principle, despite the fact that they focus on different facts about people and their actions.

The conviction standard and the serious blame principle are both grounded on the same basic intuition about justice. The intuition is really just a further specification of the intuition described earlier about the asymmetrical justifiability of harm and benefit. Justice requires arguments for harming people to be held to a higher standard than arguments which are not for harming anyone, and it requires arguments for seriously harming people to be held to an

especially high standard: there must be no room for reasonable doubt about their soundness. This holds whether the harm at issue is blame or of some other kind.

In courts of law, an argument that someone has committed a crime is typically part of a larger argument that that person is to be given a punishment which will cause serious harm such as imprisonment (or even death, in some countries). This is why it is taken to be a requirement of justice to hold arguments that someone has committed a criminal act to the "reasonable doubt" standard. When a claim about free will serves as a premise in a justification for serious blame, it must be held to the same standard, for the same reason. That is, in this context, the claim that someone has free will plays a role in an argument for serious harm, just as the claim that someone has committed a crime typically does. For this reason, it must be held to a standard based on reasonable doubt, just as the claim that someone has committed a crime must be. So the serious blame principle has the same justification as the criminal conviction standard.

#### 4. Costless Praise and Reasonable Doubt

In this section, it will be argued that condition (iii) is satisfied for costless praise so long as it can be reasonably doubted that the candidate did not act freely. Call this the *costless praise principle*. The costless praise principle does not imply an obligation in the way the serious blame principle does. The costless praise principle merely says that condition (iii) is satisfied for costless praise if it can be reasonably doubted that the candidate did not act freely. If condition (iii) is satisfied for costless praise, along with whatever other conditions there may be, then costless praise is justified. But this does not imply that costless praise is obligatory under these conditions.

The argument for the costless praise principle must proceed quite differently than the argument for the serious blame principle. This is because there is nothing in jurisprudence that might play a role parallel to the role played by the criminal conviction standard in the argument for the serious blame principle. So this argument must proceed on the basis of more general ethical considerations.

As a first step, it may be worth pointing out that if reasonable doubt standards govern both serious blame and costless praise, then there some symmetry within the asymmetry. Condition (iii) would have different branches in the contexts of serious blame and costless praise, but the branches would be symmetrical, at least to a degree. Occam's razor requires that we try the simplest theory first, and symmetry is a form of simplicity. So it makes sense to try a theory that preserves some symmetry within the asymmetry.

This effort may seem to come at the price of setting a peculiarly low standard for costless praise. But recall the perspective on costless praise suggested earlier. Costless praise is a kind of costless benefit. When all parties agree that an action under consideration is a costless benefit, we usually expect no further justification for it. Costless praise cannot be supposed to be entirely typical in this regard, of course. Costless benefits as such are justified merely by the claim that they are costless benefits, so people who have no reason to doubt this claim have no reason to ask for any further justification. But costless praise must be seen to rest on a further claim, that is, a claim that the candidate for praise is morally responsible for the action at issue. If one doubts this further claim, it makes sense to ask for a defense of it, and part of the defense must involve giving some reason to suppose that the candidate acted freely.

But when it is asked how high a justificatory standard to apply to the claim that the candidate had free will, it makes sense to look at the moral context in which the question is

asked. It seems important that the context is an attempt to provide someone with a kind of costless benefit. If the primary purpose of morality is to get people to benefit and refrain from harming each other, then the context is an attempt to do something that helps fulfill the primary purpose of morality. So it seems unreasonable to impose a high standard. If the costless praise principle is correct, then the standard is very low, i.e. that the falsity of this claim's negation can be reasonably doubted.

There are, however, many degrees between high and very low. Why should we pick a very low standard? Consider the following scenario. Suppose a person with a neurological disorder that causes intermittent problems with intention formation and motor control knocks someone else out of the path of an oncoming truck and saves his life. Suppose it is not clear whether or not he acted intentionally, and there is no way to settle the question. (The disorder also sometimes causes confusion and memory difficulties, so we cannot take our potential hero to be authoritative on this point.) But he would not object to being given the benefit of the doubt and praised as an actual hero. He has few opportunities to be praised, so it would bring him no small happiness to be praised in this case. Suppose that there is no reason to doubt that he satisfied whatever other conditions of praiseworthiness there may be in addition to the control condition. Finally, suppose that praising him would be costless—no third parties would feel slighted, and praising him would not deprive us of an opportunity to do something else worthwhile.

Even if we put metaphysical worries about free will aside, there are good reasons to doubt that the potential hero acted freely, since even free will theorists who disdain metaphysics agree that one cannot have free will with respect to anything that is not an intentional action. Yet there seems to be a strong intuition in favor of the view that costless praise of the potential hero

as an actual hero is justified, because it is not clear that he did not act with free will, and he deserves the benefit of the doubt.

Next consider whether this intuition would change if we had more detailed information about the probability that he acted freely. Suppose that we could know that there was a 10% chance that he acted freely. Would we be inclined to praise him then? I think so. How about a 1% chance? Since the praise is costless, and the potential hero benefits from being treated as an actual hero, it is not clear that anything could give us a reason to accept 10% as good enough, but not 1%. This line of thought pushes the standard downward until we reach the reasonable doubt threshold. No reasonable person could justifiably praise him if it cannot be reasonably doubted that he did not act freely. But since the praise is costless, it would seem that a reasonable person could justifiably praise him so long as it could be reasonably doubted that he did not act freely. If this is right, then the costless praise principle is correct.<sup>10</sup>

It might be objected that, by focusing on the valuable experience that the potential hero will get from being praised as an actual hero, this scenario obscures the fact that the valuable experience cannot on its own be a sufficient reason for praise, since praise is only justifiable when agents deserve it based on their actions. But this objection misses the point that there *is* an action-based reason for praise in this case, i.e. that it can be reasonably doubted that the potential hero did not act freely. This is a weak reason, but it is still a reason. The idea is that a weak action-based reason is strong enough, given that the context is an attempt to give a costless benefit to someone who deserves to be given the benefit of the doubt.<sup>11</sup>

This argument may amount to something less than a conclusive defense of the costless praise principle. But I hope enough has been said to show that it is worthy of consideration. There may be other issues relevant to free will theory where reasonable doubt standards similar

to the one in the costless praise principle play a role. Consider the "ought implies can" principle. This principle implies that I should only believe I ought to do something if I have reason to believe that I can do it. But how strong does this reason have to be? Do I need to establish that I cannot reasonably doubt that I can do it? Presumably not. I can reasonably doubt that my car will start tomorrow, and if my car does not start I cannot get to campus to teach my class, but presumably this does not conflict with my belief that I ought to teach my class tomorrow. Do I need to establish that I have better than a 50% probability of being able to teach my class to believe that I ought to do it? This standard also seems too high. Perhaps all that is required is that it is possible to reasonably doubt that I cannot do it.<sup>12</sup>

In the next section, it will be argued that no matter what theory of free will we hold, if we think that a debate about whether someone acted freely is philosophically valuable, we must accept that it can be reasonably doubted both that he did act freely, and also that he did not. If this is right, and if the serious blame and costless praise principles are correct, then it follows that anyone who thinks a debate about whether someone acted freely is philosophically valuable must accept that he does not satisfy condition (iii) for serious blame, and that he does satisfy it for costless praise.

### 5. Reasonable Doubt About Free Will

In this section, it will be argued that no matter what theory of free will one holds, if one thinks that a debate about whether someone acted freely is philosophically valuable, one must accept that it can be reasonably doubted both that he did act freely, and also that he did not.<sup>13</sup> It may be helpful to begin with a few terms for describing debates. Some debates can be represented as focused on a central claim. A debate about whether someone acted freely is an

example. Its central claim is that this person (i.e. a person of such-and-such a description) acted freely in this situation (i.e. a situation of such-and-such a description). Debates which are focused on a central claim include a pro side and a con side. The pro and the con side each have a basic position. The basic position of the pro side is that the central claim is true. The basic position of the con side is that the central claim is false.

As these terms are understood here, one can be on the pro or con side of a debate without believing it to be philosophically valuable. One person might hold that the central claim is true, and another might hold that it is false, but they might agree that the debate about it is a waste of time. The first of these people would be on the pro side, and the second would be on the con side, but neither would see any value in the debate. Someone of a different temperament might take a side and might value the debate, but as bracing intellectual exercise and nothing more. To believe a debate to be philosophically valuable is to believe it to be truth-conducive. This is to believe that working on at least some of the objections to one's basic position posed by the opposite side either tends to prompt one to revise the theory one uses to support one's basic position in a way that makes it more likely to be true, or helps one to better understand why one's existing theory is true.

Once the link between philosophically valuing a debate and believing it to be truth-conducive is established, the rest of the argument is quite short. Objections can only be truth-conducive if they are reasonable, and reasonable objections are grounds for reasonable doubts about the basic positions to which they are objections. This means that, whether one is on the pro or con side of a debate, if one takes the debate to be philosophically valuable, one must accept that it is possible to reasonably doubt one's basic position. So, in the case of a debate about whether someone acted freely, those on the pro side must accept that it is possible to

reasonably doubt the claim that he acted freely, and those on the con side must accept that it is possible to reasonably doubt the claim that he did not act freely.

Two points of clarification may be helpful before the implications of this argument for free will are discussed in more detail. First, I am not claiming that one can only get closer to a true theory, or gain a better understanding of why one's existing theory is true, when one is working on reasonable objections. One can of course be struck by a good idea at any time, even when one is working on unreasonable objections. The claim intended here is the weaker claim that, since unreasonable objections do not direct one's attention to features of one's theory which are reasonably thought of as implausible, there is nothing about unreasonable objections *as such* which would justify one in believing that working on them would be truth-conducive. One might suppose that they could haphazardly cause one to get closer to a true theory, or to a better understanding of why one's existing theory is true, but not that they would *lead* one to these outcomes.

Second, I am not claiming that everyone who philosophically values a debate must actually be able to doubt his own basic position in the debate. It is not rare for people to become so deeply committed to their basic positions that it becomes psychologically impossible for them to doubt them. But this is no objection to this argument, because the fact that it is psychologically impossible for some people to doubt a claim does not imply that it cannot reasonably be doubted. I do not even claim that *if* the debaters were reasonable, *then* they would doubt their basic positions. Philosophical disagreement is a complicated matter. As far as this argument is concerned, there may be what Richard Feldman calls "mutually recognized reasonable disagreement" between the pro and con side. That is, it may be that even when both sides recognize that the other side can reasonably doubt their basic position, they can remain

reasonable without doubting their own basic position.<sup>14</sup> To defend this argument, it is enough to claim that those on both sides would not be unreasonable if they came to doubt their own basic positions, and it seems fair to claim this much.<sup>15</sup>

Philosophers will disagree about when it is philosophically valuable to debate whether someone acted freely. The cases most congenial to free will are ones with normal adults acting in normal conditions—call such cases "normal cases". Strawsonians of a certain stripe might think the belief that people often act freely in normal cases is so fundamental to the meaning of discussions about free will that those discussions lose all sense if we put it in question. Such Strawsonians might therefore hold that debating whether people in normal cases ever act freely is not truth-conducive. But this is a minority view. Others may be committed to varieties of reductive physicalism which make the idea of free will look too absurd for a debate about it to be philosophically valuable even in normal cases. But this too is a minority view.

Anyone who thinks it is philosophically valuable to debate whether people in normal cases act freely must accept that it can be reasonably doubted that they do. If it is right that normal cases are the situations where free will is most likely to be found, then it seems to follow that if it can be reasonably doubted that people act freely in normal cases, then it can be reasonably doubted that anyone ever acts freely. If the serious blame principle is correct, this implies that anyone who believes this debate to be philosophically valuable must accept that condition (iii) is never satisfied for serious blame, and therefore that serious blame is never justified.

Anyone who thinks it is philosophically valuable to debate whether people act freely in normal cases must also accept that it can be reasonably doubted that they do not act freely. If the costless praise principle is correct, this implies that anyone who philosophically values this

debate must accept that people in normal cases satisfy condition (iii) in the context of costless praise. There are probably non-normal cases which also fit this bill. Non-normal cases do not provide ideal conditions for finding free will, but even non-ideal conditions may sometimes provide room for a philosophically valuable debate. There are lots of cases where it is hard to see room for a philosophically valuable debate about free will, for example, cases involving people who do things by accident or because of irresistible compulsions, and cases involving very young children or people with very profound cognitive or behavioural disabilities. But there may be room for philosophically valuable debate in cases involving people who do things because of apparently resistible compulsions, and cases involving older children or people with less profound cognitive or behavioral disabilities. The argument of this section implies that if we think debates are philosophically valuable in these cases, then we must accept that condition (iii) is satisfied for costless praise in these cases as well.

#### 6. Another Comparison with Wolf

The theory sketched in this paper is similar to Wolf's asymmetry theory in the sense that both hold that there is a higher justificatory standard for blame than for praise. But Wolf's theory does not rule out serious blame in the way the present theory does. As mentioned earlier, on Wolf's theory, there is an asymmetry in the control condition of moral responsibility in the contexts of praise and blame, such that alternative possibilities are necessary for blame but not for praise. But she does not hold that no one has alternative possibilities. She holds that only some kinds of determinism are incompatible with alternative possibilities. She holds that if someone is psychologically determined to act unethically (by an abusive upbringing, for example) then she could not have done otherwise. But she is a compatibilist about physical

determinism and alternative possibilities. On her view, the sheer fact that there is a deterministic physical explanation of someone's unethical action does not imply that she could not have done otherwise. She holds that there may be people who are physically but not psychologically determined to act unethically. She does not explain the distinction between psychological and physical determinism in as much detail as might be desired.<sup>16</sup> But its upshot for her theory is that some people may be blameworthy, and her theory implies no limit to how seriously they can be blamed. So her theory leaves open the possibility that it is justifiable to seriously blame some people. In contrast, the argument I have made here is silent about whether anyone may be blameworthy. But it implies that if one thinks the debate about whether anyone acts freely is philosophically valuable, then one must accept that serious blame is never justified. Wolf's view and mine seem to be mutually consistent, however.

### Conclusion

To conclude, I would like to emphasize a point which is implicit in the above discussion, but which may be worth repeating in order to make it explicit: it seems possible for free will believers and free will skeptics to accept the line of argument presented here without giving up their basic positions. By "free will skeptics", I mean those who hold that the control condition of moral responsibility is never satisfied, and by "free will believers", I mean those who hold that it is sometimes satisfied. It seems consistent to hold both that some claim is true and that it can be reasonably doubted. If this is right, free will skeptics could accept that it can be reasonably doubted that candidates for costless praise do not act freely, and if they accept the costless praise principle, they should conclude that costless praise is justifiable (at least in cases where there is no reason to doubt that the other conditions are satisfied). Similarly, believers could accept that

it can reasonably be doubted that targets of serious blame act freely, and if they accept the serious blame principle, they should conclude that serious blame is never justified. This would allow skeptics and believers to preserve their basic positions while meeting each other halfway on some important issues.

## NOTES

<sup>1</sup> She does not speak in exactly these terms, but I take this to be one of her points. In "Asymmetrical Freedom" (*Journal of Philosophy* 77.3, 1980), for example, she writes the following:

we have stronger reasons for wanting acts of blame to be justified [than we have for wanting praise to be justified]. If we blame someone or punish him, we are likely to be causing him some pain. But if we praise someone or reward him, we will probably only add to his pleasures. To blame someone undeservedly is, in any case, to do him an injustice. Whereas to praise someone undeservedly is apt to be just a harmless mistake. (p. 156)

I take it that Wolf's claim that we have stronger reasons for wanting acts of blame to be justified involves the idea that acts of blame require a stronger justification than acts of praise.

<sup>2</sup> These terms are used in a broad sense meant to cover all the various accounts of the control condition of moral responsibility, including strong libertarian notions such as dual control, and weaker compatibilist notions such as guidance control or hierarchical control. For guidance control, see John Martin Fischer and Mark Ravizza, *Responsibility and Control: a Theory of Moral Responsibility* (Cambridge: Cambridge University Press, 1998). For hierarchical control,

see Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68.1 (1971).

<sup>3</sup> Frankfurt holds that alternative possibilities are not necessary for moral responsibility. See "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy* 66 (1969), 829-839.

<sup>4</sup> Some might object to this distinction between action- and personhood-based desert. For example, it is common among free will theorists to assume that actions are the only desert bases. It is by no means absurd to suppose that actions are the only kind of desert base, since the category of action-based desert claims is very broad. Desert claims about praise, blame, reward, and punishment are all typically based on actions. Action-based desert claims also include claims that might initially appear to be character-based, because fair-minded ethicists recognize that legitimate desert claims can be based on people's character-traits only if those people have acted in such a way as to develop or cultivate those traits. Yet despite the breadth of the category of action-based desert claims, there are desert claims that cannot plausibly be supposed to be based on actions. For example, a person deserves respect, access to her rights, equal treatment before the law, not to be used as a mere means to others' ends, and to be given the benefit of the doubt, and she deserves these things not because of facts about her actions, but simply because she is a person. In other words, these kinds of desert are based not on actions, but on personhood. I discuss the role of personhood-based desert in free will theory in more detail in [author's paper 1].

<sup>5</sup> In my view, not all remorse is based on self-blame. In [author's paper 2], I argue that there is a kind of remorse which is based on suffering in sympathy with the person one has harmed rather than on self-blame.

<sup>6</sup> Punishment does not have to be retributively justified, and the argument of this paper is only relevant for retributive justifications of punishment. For a recent alternative, see Pereboom's quarantine justification in his book *Living Without Free Will* (Cambridge: Cambridge University Press, 2001).

<sup>7</sup> Egalitarian praise may not be costless when people who achieve more than others protest that they deserve to be singled out for praise in a way that excludes the others. I do not want to take a position on the legitimacy of this protest (though I am suspicious of the claim that egalitarian praise is ever unjust, and I am inclined to see sorrow caused by not being included in praise as more ethically important than sorrow caused by having to share praise with others). For present purposes, I only wish to claim that, in cases where people who achieve more than others choose not to protest egalitarian praise (perhaps out of magnanimity), they have not been treated unjustly, and there is no obstacle here to costless praise. (Also see note 10 below.)

<sup>8</sup> We might instead think that our beliefs are open to moral critique even when they are not expressed in action, and that we ought to hold our beliefs about moral responsibility to a higher standard in the context of blameworthiness than in the context of praiseworthiness. But if ought implies can, then this implies epistemic voluntarism, and this is controversial.

<sup>9</sup> The argument of this section is based on a longer, more detailed argument which I present in [author's paper 3].

<sup>10</sup> Suppose that someone else, who has no such disorder, was praised for rescuing a different person from certain death last week. He might be pained if the potential hero got the same kind of praise he got—he might object that there is more reason to doubt that the potential hero acted freely than there is to doubt that he acted freely, so it is unfair to give the potential hero the same

kind of praise he got. I have my doubts that this would be unfair. But as far as the argument of this paper is concerned, I am willing to concede that praise would not be costless in this case. It is enough to point out that last week's rescuer might be magnanimous enough not to be pained or to object, and that in this case, it is possible to give our potential hero the same kind of praise last week's rescuer got without being unfair. (Also see note 7 above.)

<sup>11</sup> Recall that, according to the view taken here, we all deserve to be given the benefit of the doubt, not because of how we have acted, but simply because we are people. (See section 1, especially note 4.) It must also be emphasized that this appeal to the context need not be understood as a purely consequentialist move, since the claim is not simply that it is good for the potential hero to be given the benefit of the doubt, but that he deserves it.

<sup>12</sup> I discuss the problem of understanding the "ought implies can" principle in the context of free will skepticism in more detail in [author's paper 4].

<sup>13</sup> This argument is a more general version of an argument I describe in [author's paper 3].

<sup>14</sup> See Richard Feldman, "Epistemological Puzzles about Disagreement", in S. Hetherington (ed.) *Epistemology Futures* (Oxford: Oxford University Press, 2006) pp. 216-236. Feldman is a skeptic about mutually recognized reasonable disagreement. His argument suggests that, to be reasonable, philosophers on the pro side of the debate about whether we have free will might have to suspend judgement about the central claim when they recognized that philosophers on the con side had reasonable doubts about the central claim.

<sup>15</sup> This argument implies that anytime we think it is philosophically valuable to debate a proposition, we must acknowledge that it can be reasonably doubted, and some might object to this. For example, it implies that anyone who thinks that the debate about the existence of other

minds is philosophically valuable must accept that it is possible to reasonably doubt that other minds exist, but it is clearly absurd to claim that just because someone thinks this debate is philosophically valuable, she should act as if the existence of other minds has been put in doubt. In my view, it is true that anyone who thinks the debate about other minds skepticism is philosophically valuable must acknowledge that the existence of other minds can be reasonably doubted. But reasonable doubt about the existence of other minds does not imply that we should act as if other minds do not exist. It would only imply this if we accepted a principle which said that reasonable doubt about the existence of other minds requires us to act as if other minds do not exist, and I cannot see any reason to accept such a principle. This highlights an important difference between doubts about free will and doubts about other minds. If the argument of this paper is right, morality requires us to treat reasonable doubt about free will as practically significant. But morality could not require us to treat reasonable doubt about the existence of other minds as practically significant: if the primary purpose of morality has something to do with getting people to benefit one another, or at least getting them to refrain from harming one another, then morality requires a commitment to the existence of other minds.

<sup>16</sup> For a critique of this distinction, see Gary Watson's review of *Freedom Within Reason*, *Philosophical Review* 101.4 (1992), pp. 890-893.