



VICTOR TADROS 

## Distributing Responsibility

A widespread view in moral, legal, and political philosophy, as well as in public discourse, is that responsibility makes a difference to the fair allocation or distribution of things that are valuable or disvaluable independently of responsibility. For example, the fairness of punishing a person for wrongdoing varies with her responsibility for wrongdoing; the fairness of requiring a person to pay compensation varies with her responsibility for the harm that she caused; the fairness of one person being worse off than another varies with her responsibility for being worse off; the fairness of inflicting defensive harm on a person to avert a threat varies with her responsibility for causing or posing the threat; and so on.

Little attention has been paid to the central issue of this article: the allocation and distribution of responsibility itself. How can responsibility be allocated or distributed? The social structures of a society, and the choices that individuals make within them, make a difference to who will be responsible for what and how responsible they will be. A person's responsibility for wrongful actions, imprudent actions, prudent actions, good actions, supererogatory actions, and so on, is itself influenced by social structures and choices. Given their impact on what people will be responsible for, how should these social structures be developed, and choices be made?

The allocation and distribution of responsibility can be fair or unfair: responsibility for conduct itself, I argue, makes decisions just or unjust, and that affects the just distribution of other things, such as welfare-generating resources. Furthermore, the potential injustice of inequalities in responsibility can be counterbalanced by reverse inequalities in welfare-generating resources. This second idea has radical implications for the

I am grateful to audiences at Warwick, Oxford, and Newcastle. Thanks, for useful discussion and comments, to Matthew Clayton, Helen Frowe, Alex Kaiserman, Michael McKenna, Dana Nelkin, Tom Parr, Derk Pereboom, Massimo Renzo, Adam Slavny, Patrick Tomlin, Andrew Williams, and Steven Wu, as well as two Associate Editors of *Philosophy and Public Affairs*.

© 2020 The Author. *Philosophy & Public Affairs* published by Wiley Periodicals, Inc. *Philosophy & Public Affairs* 9999, no. 9999

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

relationship between welfare or resources on the one hand and responsibility on the other—in distributive justice, there is at least some pressure to allocate welfare or resources to those who are responsible for wrongdoing, and away from those who are responsible for good deeds.

### I. What Is responsibility?

Grasping the idea that responsibility can be allocated or distributed involves understanding responsibility. I am concerned with one sense of responsibility—the sense involved in attributing actions to people. This sense of responsibility is invoked in propositions like: Mark Chapman was responsible for killing John Lennon; Jane Austen was responsible for writing *Pride and Prejudice*; Marie Curie was responsible for discovering the theory of radioactivity; and so on.

I don't have space for a detailed defense of any particular view of responsibility. However, I will make some assumptions about it. I assume compatibilism: that responsibility is compatible with determinism. I also assume that people are responsible for their conduct in the wide range of circumstances that roughly corresponds to the folk view. Absent exemptions or excuses, people are normally responsible for what they do intentionally, voluntarily, recklessly, negligently, and so on.

One reason for these assumptions is that I have compatibilist sympathies, as do most people who write about responsibility, including those who are concerned with the role of responsibility in matters of justice. And most compatibilists, including me, don't have extremely revisionist views about who is responsible for what in the real world. Furthermore, compatibilism helps us to brightly illuminate the distributive problems that I will raise. In Section IX, I briefly explore the question I am concerned with for libertarians.

Here is a more precise and complete statement of my assumptions. A person is responsible for her conduct because her acts are caused by certain agential facts—what I will just call “the relevant agential facts.” Compatibilists disagree about these facts. Some think that people are responsible for acts because they issue from the correspondence of first- and second-order preferences or desires; others because their acts issue from what they value, or from their character; others because they were able to do otherwise (in some sense compatible with determinism); and

others because their actions were governed by suitably reason-responsive mechanisms. I remain neutral on that issue.

I also assume that some standard responses to a person because she was responsible for her conduct are apt, though not that any particular response is warranted. The range of responses include certain conversational practices about a person's wrongdoing, certain emotional responses to the person such as resentment, anger, indignation, sadness, and some kinds of blame. I assume a similar thing about responsibility for intentional good acts—that gratitude and praise, for example, are apt because the person was responsible for her good acts.<sup>1</sup>

I do not assume any normative or axiological relationship between welfare and responsibility: that it is better, or more just, that those who are responsible for wrongdoing are (or are made) worse off, or that those who are responsible for good deeds are (or are made) better off. One of my aims is to explore those issues.<sup>2</sup>

Finally, I assume that this fairly standard understanding of determinism is true:

*Determinism:* Determinism is true in some world if the complete state of that world at any time,  $t1$ , and the physical laws that apply to it, makes the complete state of that world at all later times inevitable.

As how a person acts is part of the complete state of the world at the time at which they act, compatibilism is thus understood as the view that a person can be responsible for what she has done even though it was

1. Though the article is cryptic, Peter F. Strawson's essay, "Freedom and Resentment," in *Freedom and Resentment and Other Essays* (London: Methuen, 1974), did a great deal to stimulate interest in this idea. Prominent defenders of this approach include Gary Watson, *Agency and Answerability: Selected Essays* (Oxford: OUP, 2004); Thomas M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard UP, 1998), chap. 6; Thomas M. Scanlon, *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, Mass.: Harvard UP, 2008), chap. 4. Even some who are skeptical that we can deserve blame, because we lack the free will needed, think that some practices like these can be apt. See, for example, Derk Pereboom, "A Notion of Moral Responsibility Immune from the Threat from Causal Determination," in *The Nature of Moral Responsibility: New Essays*, eds. R. Clarke, M. McKenna, and A. M. Smith (Oxford: OUP, 2015).

2. Some think that there is a sense or kind of responsibility that involves a person deserving to be better or worse off as a result of the relevant agential facts. I doubt that this is right. For discussion, see Victor Tadros, "Responsibility as Personal Significance" (unpublished manuscript). If I am wrong, I do not assume that we have this kind of responsibility.

inevitable that she acted as she did given the state of the world prior to her birth, and the physical laws that apply to it.

## II. The Social Dependence of Responsibility in the Real World

If these assumptions are true, responsibility is obviously socially dependent. Indeed, it is socially dependent on any plausible view of responsibility that is not radically revisionist. How many people are responsible for wrongdoing, and who these people are, depends on geography, architecture, resource distribution, parenting, education, employment opportunities, and so on. These things either involve social and political decisions, or their allocation and distribution stem from social and political decisions.

For example, the rate of violent wrongdoing is much higher in poor urban communities than in wealthy rural villages. Social and political decisions determine how many people grow up in which communities and what they are like. A different set of decisions by politicians, urban planners, and voters would have exacerbated or ameliorated criminogenic differences in social circumstances that people grow up in, resulting in different rates and distributions of wrongdoing. Only radical skeptics about responsibility conclude that the social determinants of wrongdoing undermine responsibility completely, so these social and political decisions, in conjunction with other facts, determine the rate and distribution of responsibility for wrongdoing.

Cases of large-scale social planning involve tricky nonidentity issues—which people exist, and then act wrongly, depends on large-scale social and political decisions about wealth, geography, and architecture. I limit my focus to cases where identity is fixed. Social and political decisions obviously affect who is responsible for wrongdoing in such cases too. Here is a striking example. As Elizabeth Hinton shows, two distinct approaches were used to deal with young offenders in the United States in the 1970s.<sup>3</sup> A more rehabilitative approach tended to be used for white young offenders, while a more punitive approach tended to be used for blacks. There is significant evidence that the more rehabilitative approach resulted in less recidivism and escalation of criminal activity, partly

3. See Elizabeth Hinton, *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America* (Cambridge, Mass.: Harvard UP, 2016), chap. 6.

because the punitive approach involved incarceration, which created communities of offenders. This may well have resulted in black young offenders in the 1970s committing more crimes than white young offenders later in life.

Again, only a radical skeptic about responsibility concludes that these black people are not responsible for their wrongful conduct later in life because of the decision to incarcerate them rather than rehabilitate them when they were young. Yet the racist decisions to incarcerate blacks rather than whites were obviously unjust. Part of the injustice is that incarceration increased the chance that these people would act wrongly later in life. As young black offenders had a powerful interest in not acting wrongly later in life, they had an interest in being rehabilitated rather than incarcerated and that made the decisions unjust. Or so I will argue.

Similar examples are widespread and familiar. Social policies are often designed to prevent wrongdoing or are criticized for failing to do so. Erosion of educational and social facilities for young people are criticized because they make a difference to whether young people offend; transitional processes, practices, and institutions for military personnel, as well as those who have been incarcerated, are needed in part because these people are especially likely to offend or reoffend; urban environments and school buildings need to be restored and protected, because erosion of those environments causes crime; and so on. The interests of potential victims of wrongdoing are relevant to these decisions, but so are the interests of potential wrongdoers in being guided away from wrongdoing.

Social policies also aim to ensure that people act well. Education and parenting can foster a sense of justice and are criticized if they fail to do so. Noble deeds and valuable discoveries are celebrated to inspire people to emulate and go beyond the achievements of others. The distribution of educational resources, support to parents, and accolades determines the rate of valuable acts, and who will perform them. But people are nevertheless responsible for the good deeds that result from these social policies.

Policies, practices, and decisions that determine the rate and distribution of responsibility create conflicts of interest, because they require resources. Where those resources are scarce, decisions must be taken about their distribution. How much should be spent on preventing wrongful conduct, and how much on supporting good conduct? And what principles make the distribution of resources to bring about responsibility for

good and bad acts just? These questions have received little attention in debates about distributive justice. And these familiar social questions are themselves particular instances of a much more general problem: given that responsibility for good and bad acts is itself caused by social and political decisions, how should responsibility be distributed, and how does its distribution interact with the distribution of other things that are a matter of justice?

### III. The Value and Disvalue of Responsibility

The examples just outlined have many distracting features that make it harder to focus on the distribution of responsibility as such. The decision to incarcerate young black people rather than whites, for example, is unjust independently of its effects on responsibility for wrongdoing.

In what follows I reflect on cleaner cases to show four things. First, in this section, I show that it is disvaluable for a person to be responsible for wrongdoing both for instrumental reasons and as such. This helps to establish that responsibility is part of the currency of justice, because justice is mainly concerned with the allocation or distribution of things that matter independently of justice. Second, in Section IV, I show that familiar issues in distributive justice apply to responsibility itself. Third, in Section V, I show that what is just overall depends on intrapersonal and interpersonal tradeoffs between different interests relevant to justice. And that supports my fourth conclusion: the initially strongly counterintuitive view that there are reasons of distributive justice to provide more welfare-generating resources to those who are responsible for wrongdoing than those who are not other things equal. I then explore the relationship between distributive justice and other normative considerations to indicate the complexity of the relationship between the distributive considerations considered here and other features of justice. The argument in Section VI suggests that similar conclusions to those defended on compatibilist assumptions are true if incompatibilism about responsibility is true.

#### *A. Cleaning Up*

I begin with a case that focuses our attention on the allocation of responsibility for wrongdoing, where who is responsible for wrongdoing depends on a social decision taken before a person was responsible,

and there is no reason to care about the decision independently of that fact:

*Seating Arrangement:* Billy and Bobby start a new school at age 7, and are very similar. On their first day, Teacher sits them in two free seats in class. In World 1, where Teacher sits Billy next to Jack, Jack becomes Billy's best friend. In a nearby world, World 2, Teacher sits Billy next to John and John becomes Billy's best friend. Things are *vice versa* for Bobby. In World 1, Jack is a bad influence on Billy. When they are 20, he persuades Billy to commit a single crime—a serious assault on Jeff—which Billy does intentionally and without excuse. This does not happen in World 2: Billy does not commit any serious wrongs in that world, as John is a good influence on Billy. Again, things are *vice versa* for Bobby. Other things are equal.

Recall that we are assuming *Determinism*. Let us also assume that the difference between World 1 and World 2 at the time that Billy and Bobby are seated just arises from the decision that Teacher makes. Future differences between these worlds depend on Teacher's decision. The world where Teacher decides to sit Billy next to Jack is World 1. It is inevitable in that world, at the moment that he is seated, that Billy will assault Jeff, and Bobby will not. But, obviously, Billy is not responsible, at the moment that he sits next to Jack, for anything that he does in the future. He is not morally responsible for anything at the age of 7. The world where Teacher decides to sit Billy next to John is World 2. It is inevitable in that world, at the moment that he is seated, that Bobby will assault Jeff, and Billy will not.

Suppose further that the difference between World 1 and World 2 is the most minimal thing that gives rise to Teacher making the different decisions that she makes in World 1 and World 2 (say, for example, that her head is slightly turned to Billy when the boys enter the classroom in World 1 where it is slightly turned to Bobby in World 2). That thing has no moral salience beyond its effect on her decision.

Given the assumptions outlined in Section I, although Billy's assault on Jeff counterfactually depends on Teacher's decision, and Billy is not responsible for that, he is responsible for assaulting Jeff. All compatibilists should be comfortable with this conclusion; the case is much less

controversial than cases that have been used to challenge compatibilism, such as manipulation cases.<sup>4</sup>

Unlike most real-world cases, there are no victim-centered reasons that affect the value of who sits where—Jeff will suffer the same fate whoever sits next to Jack. This helps to focus our attention on the interests of Billy and Bobby in responsibility without being distracted by the difference that responsibility for wrongdoing makes for others.

### *B. Responsibility for Wrongdoing Is Worse for Us*

In this subsection, I show that World 1 is less valuable for Billy than World 2. I make no claims about well-being in particular. There is a great deal of dispute about what well-being includes. On some views, responsibility for wrongdoing is a negative component of well-being. I do not defend this idea: only the idea that responsibility for wrongdoing matters for the sake of the responsible person.

Here are three reasons why his responsibility for wrongdoing matters for the sake of Billy, and thus that he has reason to value World 2 over World 1. The first is instrumental. People inflict harms and other costs on wrongdoers. They are subject to defensive, punitive, and compensatory harm, for example, and they are publicly criticized, shunned, and socially ostracized. Some of these responses may be warranted; others are not. But warranted or not, people have good reason to disvalue being responsible for wrongdoing because of these effects. These effects do not establish that responsibility fundamentally matters for distributive justice—only that it typically affects other things that fundamentally matter.

Second, it matters for the sake of a person that certain responses to them are (or are not) apt. For example, a person has reason to disvalue it being apt to blame her for her conduct, for others to resent her, and for

4. Perhaps the most influential challenge of this kind is Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford: OUP, 2014), chap. 4. Our intuitions about manipulation cases are less reliable, both because we find manipulation morally significant and because our intuitions are partly explained by the impression that the cases create that there is a difference in responsibility between the manipulators and those who are manipulated. Even with respect to those cases, many people continue to think that responsibility is not threatened. Furthermore, even those who think that it is focus on a narrow conception of responsibility—that concerned with basic desert. For further discussion, see, for example, Michael McKenna, “A Hard Line Reply to Pereboom’s Four-Case Argument,” *Philosophy and Phenomenological Research* 77 (2008): 142; Carolina Sartorio, *Causation and Free Will* (Oxford: OUP, 2016), Victor Tadros *Wrongs and Crimes* (Oxford: OUP, 2016) chap. 5.



her to feel guilty at what she has done, independently of whether anyone has these reactions. People often want to do things that are worthy of respect, and not worthy of contempt, not just because of the things that make these reactions apt, but because of the aptness of these reactions themselves. For example, a child might be motivated not to act in a way that would disappoint his parents, or to act in a way that they would approve of, even after his parents have died. But these motivations are warranted only if their reactions to his conduct would be apt.

Third, and most importantly for our purposes, it is disvaluable for a person's sake that she is responsible for wrongdoing as such, independently of the actual or likely consequences of being responsible, or what it makes apt. To see this, suppose that you are Billy's parent. Parents have stronger reason to hope that the world is valuable for their children than they have to hope that it is valuable for strangers, and typically do so. You have good reason to hope that Billy rather than Bobby will not be responsible for wrongdoing, quite independently of any further bad consequences that this will have for Billy, or the reactions that others would be justified in having to his conduct. For this reason, you hope, for Billy's sake, that you live in World 2, where Teacher sits Billy next to John.

To fully grasp the fact that responsibility for wrongdoing is significant in itself, it helps to consider a case where the negative consequences of a person being responsible for wrongdoing don't occur, such as:

*Amnesia:* As *Seating Arrangement*, except that everyone in World 1, including Billy and Bobby, thinks that Bobby has assaulted Jeff. This is because Billy and Bobby were drunk on the night in question; drunk enough that they cannot remember what happened, but not so drunk that they lacked responsibility for what they did. Furthermore, it appears to Jeff that Bobby rather than Billy assaults him, which he then reports. And the incident is caught on camera in a way that makes the culprit look more like Bobby than Billy, so there is good evidence that Bobby was the wrongdoer. All of the bad effects that anyone gets from the judgment that a person is responsible for the assault, such as the bad feelings, the punishment, and the legal obligation pay compensation, thus fall on Bobby. The feelings of guilt are not very severe, and the burdens that he suffers through punishment and compensation are not awful. In World 2, things are vice versa between Bobby and Billy.

Which world is more valuable for Billy? To test this, imagine again that you are Billy's parent. I take it that parents have powerful reasons to hope that their children will not be wrongly accused, convicted, and punished for wrongdoing. Not only is Billy worse off in these respects in World 2, he unjustly worse off. So, there are powerful reasons for Billy's parents to favor World 1, where Billy was responsible for wrongdoing, but is not accused, convicted, and punished for it.

Nevertheless, overall, were I Billy's parent, I would prefer World 2, where although Billy is accused, convicted, and punished for something he didn't do, he wasn't responsible for wrongdoing. Having asked many people this question, only a small minority prefer World 1 to World 2 for Billy's sake.<sup>5</sup> Even those who conclude that Billy's well-being is greater in World 1 than World 2 tend to think that well-being is not all that matters, and that responsibility for wrongdoing counterbalances differences in the loss of well-being and the unjust treatment that Billy faces in World 2. Furthermore, even those who think that World 1 is preferable for Billy's sake find the question difficult, which suggests that something powerful counterbalances the bad effects of being wrongly accused of wrongdoing in World 2.

Perhaps some may put this intuitive reaction down to the fact that a parent fails in her parental responsibilities if her child is a wrongdoer. But a similar thing is intuitive from the first-person perspective. Suppose that you wake up from a coma, knowing that you have been involved in a serious assault, but not knowing whether you committed the assault or were the victim of it. As long as the consequences are not awful, I would much prefer to be the victim of the assault. The vast majority of those I have asked share this view.

Some might argue that it is wrongdoing, rather than responsibility for wrongdoing, that explains our reaction. To test for this, suppose that on waking up from a coma, you know that you and one other person have assaulted a person together. Only one of you was responsible because the other was hypnotized. I hope that I was the one who was hypnotized and this view is also widely shared.

5. At several points in the article, I will indicate that certain intuitions are widely shared or supported. This is the impressionistic result of discussion with many people about the cases in seminars, lectures, and more informally. This is admittedly unscientific!

Perhaps it might be argued that these things are true because of the ways in which it is apt to respond to those who are responsible for wrongdoing. It is difficult to test for this, but when reflecting on *Amnesia*, it doesn't seem that the reason for our preferring World 2 for Billy's sake is mainly to do with the apt reaction to Billy's conduct, rather than responsibility for wrongdoing itself.

Just as responsibility for wrongdoing matters as such, so does responsibility for good deeds. Consider:

*Seating Arrangement (Good Variation):* As *Seating Arrangement* but Jack is a good influence, so that rather than assaulting Jeff, Billy but not Bobby performs a supererogatory life-saving act in World 1, whereas Bobby but not Billy does so in World 2.

Billy's parents have reason to hope that Billy is in World 1. That is also true in *Amnesia (Good Variation)* where if Billy performs the supererogatory act, Bobby gets the credit. As before, these judgments are warranted because responsibility for supererogatory acts matters as such and not just because of the consequences, or what the considerations make apt. Furthermore, as before, we can see that responsibility has value from the fact that Billy has reason to hope that if he performs the supererogatory rescue, he does so voluntarily, and not, for example, from hypnosis.

Why do people have a reason to prefer that they are responsible for good deeds, and not responsible for wrongdoing? This is a difficult question, and nothing I say depends on any particular answer. But I can sketch a suggestion. As well as their capacity for pleasure and pain, people have distinctive and significant moral value and moral status because they are able to critically reflect on values and reasons for action, and respond to them appropriately. The exercise of these abilities is in keeping with what makes the person have moral status when they are exercised well, and it is in conflict with what makes the person have moral status when they are exercised badly. Where a person acts wrongly, she either fails to critically reflect on values and reasons for action appropriately, or she fails to respond appropriately. And that gives her life less value because her conduct is in conflict with what gives her moral status. When a person acts very well, the contrary is true.

When I say that the person's life has less or more value, I don't mean that her life is less important from a moral point of view, or that people

have less reason to care about what happens to her. As we shall see, there is some sense in which the opposite is true. The importance of a person's life depends on her moral status, and not on what she does. I mean, rather, that her life is less good, and she therefore has less reason to value the fact that her life is lived in that way.

#### IV. From Value to Justice

Having established that being responsible for conduct is valuable for the sake of the responsible person (in the case of supererogatory acts) or disvaluable for the sake of the responsible person (in the case of wrongful acts) in itself, I turn from evaluative to normative questions. Because responsibility has value or disvalue for a person, and responsibility counterfactually depends on decisions of others, responsibility itself can be distributed justly or unjustly. The real-world examples offered in Section II already suggest this. But the view can be more clearly and powerfully exemplified and supported using hypothetical cases.

In the real world we are often uncertain whether our social decisions will result in particular people being responsible for wrongdoing. For example, those who incarcerated black people in the 1970s did not know that any particular person would be responsible for wrongdoing as a result of their decisions, even if they had evidence that a larger proportion of the incarcerated population than the non-incarcerated population would be.

This can create the misleading impression that future wrongdoing was the result of choices made by the incarcerated population *rather than* the decision to incarcerate them. The temptation to see things this way arises because we have the pre-theoretical intuition that if our conduct is chosen, it is not fully determined by events prior to choice. But, assuming compatibilism, the right way to understand these cases is as follows: some people's wrongdoing counterfactually depends on the decision to incarcerate them; that decision was a feature of the state of the world prior to their wrongdoing that makes their wrongdoing inevitable; and yet that is compatible with their choosing to act wrongly, and being responsible for their wrongdoing.

To avoid the distracting incompatibilist temptations, let us consider cases where it is clear that social decisions prior to a person being responsible for anything make the difference between it being inevitable that she is responsible for wrongdoing in the future and it being inevitable that she

is not. And such cases have a further virtue. There are powerful arguments that in a deterministic world there are no objective probabilities.<sup>6</sup> And even if there are, they make no difference to who is responsible for what. Furthermore, after a person is responsible for wrongdoing, we can be virtually certain that her responsibility counterfactually depended on, earlier social decisions. We may be uncertain about which decisions made the relevant difference (although we are sometimes confident about that too). But all conduct is caused by, or counterfactually depends on, prior social decisions of others; responsibility for wrongdoing is no exception.

With this in mind, consider:

*Vision: As Seating Arrangement*, except Teacher has a highly credible and true vision of the future just prior to seating Billy and Bobby. She knows exactly what will occur in their futures, and thus that who will be responsible for wrongdoing depends on whom she sits next to Jack. She also knows that immediately after seating the children, she will forget her vision, and so will be unable to influence their futures in the light of what she knew. She must seat the two children in the two free seats, or some much worse fate will befall them both.

Some may have skeptical worries about the possibility of gaining knowledge about the future through visions. I propose to leave these worries aside. The falsity of this has no bearing on my argument—if you think that visions cannot provide decisive evidence, just fill in some other mechanism that you find adequate—perhaps, for example, Teacher needs a short spell of extraordinary powers of evidence gathering and reasoning.<sup>7</sup> Or just imagine that Teacher has extremely powerful evidence about the future from some other source.

6. David Lewis, for example, thought it obvious that there are no objective chances (other than one or zero) in a deterministic world. See David Lewis, "Postscripts to 'A Subjectivist's Guide to Objective Chance'," in *Philosophical Papers Volume II*, ed. David Lewis (Oxford: OUP, 1986), pp. 118–21. For a more developed analysis, see Jonathan Schaffer, "Deterministic Chance?," *The British Journal for the Philosophy of Science* 58 (2007): 113.

7. Some think that there are in principle limitations on the ability of embedded systems in a deterministic universe to predict future conduct. If there are, though, they arise because of special cases where future systems are disposed to respond to the prediction itself. See, for example, Stefan Rummens and Stefaan E. Cuypers, "Determinism and the Paradox of Predictability," *Erkenntnis* 72 (2010): 233. As Billy and Bobby have no reason to suspect that Teacher has predicted the future, this issue is irrelevant to this case.

As it is in Billy's interests not to be seated next to Jack, Teacher has a reason not to sit him there. Of course, she has an equally weighty reason not to sit Bobby there, and it may seem that Teacher's vision makes no difference to what she should do. On the contrary, where there is a significant good or a burden to be distributed, the good or burden must be distributed through a fair process. Where only one person can receive the good or burden and other things are equal, we must ensure that each person has an equal chance of receiving that good or burden, and that it is not distributed for improper reasons.

This is no less true with respect to responsibility for wrongdoing. Given their interests, Teacher should ensure that each has a fair chance not to be seated next to Jack, and she should ensure that any bias that she has toward Billy and Bobby does not influence her decision. She could do this by flipping a coin. Coin-flips are not normally required where nothing is at stake—were everything equal between the two seats, Teacher would have no reason to flip a coin. This suggests that fairness determines what Teacher should do.

The distributive significance of responsibility for wrongdoing is even more powerfully illustrated by cases where further distributions of responsibility for wrongdoing are available. Consider:

*Vision (Three Options):* As *Vision* except Teacher has a third option: to sit Billy and Bobby next to each other. If she does this, each will commit a less serious wrong against Jeff. This will make no difference to Jeff, because suffering the two less serious wrongs is equally bad for him overall as suffering one more serious wrong, committed by either Billy or Bobby.

The intuition that Teacher ought to sit Billy and Bobby next to each other is widely shared. Some might find it better to concentrate responsibility for wrongdoing on either Billy or Bobby than to share it between them, perhaps on the basis that it is especially important to keep one person morally pure where this is possible. The more attractive view, and one that is widely shared, is similar to that we find in other cases of distributive injustice. If a cost or burden of a certain fixed magnitude must be distributed between two people, and other things are equal, it should be shared between the two, rather than being placed wholly on one person. A wide range of moral views support this idea.

This is just one of a wide range of parallels between our distributive intuitions, theories, and judgments about responsibility for wrongdoing and other things that are matters of justice. For example, familiar dilemmas in distributive justice arise in the case of responsibility for wrongdoing. And these dilemmas seem susceptible to just the same kind of analysis as the distribution of other currencies. For instance, there is a dilemma in distributive justice whether preventing a very trivial harm to a large number of people can counterbalance preventing a very significant harm to one person.<sup>8</sup> A similar thing is true in the case of responsibility for wrongdoing. Responsibility, on any sensible view, is scalar. Exactly what makes it scalar is a difficult question that I set aside. Now suppose that Teacher can either make it true either that Billy is highly responsible for wronging Jeff, or that many people bear a tiny amount of responsibility for wronging Jeff. If the overall magnitude of responsibility in the latter case is greater than in the former case, ought Teacher to make it true that Billy is highly responsible or that the second group each has a very small amount of responsibility for doing that? This seems a difficult question, and it is difficult in just the same way as the parallel case in harm distribution. I don't mean to imply, here, that increases in one person's responsibility for some event necessarily reduces another's – two people can be fully responsible for the same event, so that there is a necessarily a fixed amount of responsibility that needs allocating amongst people. I only mean to point to the possibility of distributive dilemmas.

Or consider the dilemma whether to make one person very well off and another very badly off, or to make two people moderately well off. Even if aggregate welfare or resources is higher if the two are made unequally well off, there is some pressure to make the two equally well off. The same thing is true in the case of responsibility for wrongdoing. Suppose that two people will act wrongly, but their degree of responsibility for their wrongdoing will vary, say because the extent to which they are involuntarily intoxicated is varied. Either one person will be fully responsible for the wrongdoing and the other is not responsible at all, or both will be moderately responsible for their wrongdoing. Even if the total magnitude of responsibility is smaller if it is distributed unequally, there is some pressure to make the two equally responsible.

8. The literature is vast, but perhaps the most comprehensive treatment is Larry Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Oxford: OUP, 2012).

Difficult questions about the deeper grounds of principles that explain egalitarian results in simple cases of welfare or resource distribution also apply to responsibility to wrongdoing.<sup>9</sup> For example, the intuition that responsibility for wrongdoing should be shared equally in *Vision 2* might be defended on a deontic or a telic basis. The former view is that responsibility for wrongdoing should be shared equally because that is the upshot of the duties that Teacher owes to Billy and Bobby to give their interests equal weight in her practical reasoning. The latter view is that an equal distribution of responsibility for wrongdoing is more valuable (or less disvaluable) than an unequal distribution, and Teacher has good reason to bring the most valuable distribution about. It might also be defended on an egalitarian or prioritarian basis. The former view is an essentially comparative view: that equality of responsibility for wrongdoing is itself morally significant. The latter is not essentially comparative: the justice of an equal distribution in simple cases is just the upshot of the fact that increments in responsibility for wrongdoing have greater moral significance where a person is more responsible for wrongdoing than where a person is less responsible for wrongdoing.

Though I cannot answer difficult questions about the right distributive principles, or their grounds, here, there is at least some initial reason to suppose that there are some general answers to these questions that apply to welfare or resources on the one hand and responsibility for wrongdoing on the other. A general project in distributive justice, then, is to consider which questions apply equally to different justice relevant interests, and which do not.

#### V. Responsibility, Opportunity, and Autonomy

Here I consider two objections to the view that responsibility for wrongdoing is a matter of distributive justice.

##### A. *Opportunity, Influence, and Actuality*

First, it might be argued that the cases that I have outlined do not show that responsibility for wrongdoing itself, but rather that the opportunity

9. Perhaps the most significant clarification of these disputes is Derek Parfit, "Equality and Priority?," in *The Ideal of Equality*, eds. M. Clayton and A. Williams (Basingstoke: Palgrave, 2002).



not to be responsible for wrongdoing, is a matter of distributive justice. In *Seating Arrangement* Billy's opportunity not to act wrongly is worse than Bobby's in World 1, it might be argued, because of Jack's bad influence. Teacher's decisions are just or unjust because of the distribution of that opportunity, not the distribution of responsibility for wrongdoing itself.

One natural understanding of an opportunity is that opportunities are shaped by the actual or expected results that one's conduct can have, but not by influences over one's conduct. On that view, credible threats typically alter opportunities, because they add expected or actual costs to one's decisions. But pressure and persuasion need not—they may simply alter the likelihood that one will act, but not the expected or actual results of one's actions without affecting the value of what is done. On that view, assuming that people are able to do otherwise than they did, Billy and Bobby do not have importantly different opportunities.

Perhaps it might be argued that acting wrongly comes at a price for Billy in World 1—he will seem less cool in Jack's eyes, and that will erode a friendship that he values. Let us, then, make this fact equal between Billy and Bobby in each world:

*Vision (Tone Variation):* As *Vision* except John will also encourage whoever sits next to him to act wrongly, but that person will decide not to do so. This will be so because Jack and John have different voice tones, and due to their particular and relevantly similar genetic makeup, Billy and Bobby each find Jack's voice tone more persuasive. For that reason, Billy assaults Jeff in World 1, where Bobby refrains from committing a similar assault, and his friendship with Bobby is eroded. Things are reversed between Billy and Bobby in World 2.

In *Vision (Tone Variation)* we have done a great deal to ensure that the opportunities of Billy and Bobby are identical; it is the outcome of their having these opportunities that differs. But the case is morally identical to *Vision*.

This is not to say that opportunities lack moral significance independently of what people choose. There may be a range of reasons to value unselected options. For example, the value of what we choose may depend on whether we choose between valuable options, and we may shape our own lives only if we decide between available options. But whatever the selection-independent value of options, we also value, or disvalue, responsibility for conduct as such, as we have seen.

Perhaps the critic might then suggest that what matters for distributive justice is not just the distribution of opportunities, as I have understood them, but the distribution of both opportunities and things that might causally influence conduct.<sup>10</sup> That would explain why *Vision (Tone Variation)* is morally identical to (*Vision*). But there are no reasons of justice to distribute things that might causally influence conduct independent of their actually doing so, unless those things have some further valuable property. And they need not. To see this, suppose that in *Vision (Tone Variation)*, whoever sits next to Jack will find Jack's voice more persuasive, but this person will nevertheless not act wrongly because he will be deterred from doing so, and all else is equal between Billy and Bobby. Now it hardly matters who sits next to Jack.

Perhaps the critic might press on and claim that it is the distribution of things that actually cause responsibility for wrongdoing that are the subject of distributive justice rather than responsibility for wrongdoing itself. This view is extremely hard to believe. The causes of responsibility for wrongdoing need have no moral significance at all independent of what they cause. It is hard to see why the distribution of the causes could be a nonderivative currency of justice. It is the distribution of responsibility for wrongdoing itself that we care about.

### B. *Respect for Autonomy*

A second objection is that respect for autonomy rules out responsibility for wrongdoing a matter of distributive justice. One way to understand this idea is that making decisions about responsibility for wrongdoing on the basis of the interests of those who will be responsible is problematically paternalistic.

While this may be true of some cases, though, it is not true in the *Vision* cases. For example, Teacher's decision not to seat anyone next to Jack in *Vision 2* is not objectionable on the grounds that this problematically interferes with the autonomy of the boys for their own sake. There

10. Some might think that things that influence, or tend to influence, us affect our opportunities, where others might think that they do not, but that distributive justice is concerned with them, as well as with opportunities. This is just a terminological choice that makes no substantive difference.

are many reasons why. First, Teacher's decision does not prevent the boys from exercising their autonomy; rather it determines how they will exercise their autonomy. And, at the moment the decision is taken, they are young enough that paternalistic interference is unproblematic. Neither Billy nor Bobby is controlled against their will for their sakes at a time when they are autonomous agents.

Second, not all methods of shaping the lives of others for their own sake involve the kind of interference that is problematically paternalistic. At one end of the spectrum, one person can affect another's decisions by engaging with her autonomous agency—by giving advice or making legitimate moral demands of others, for example. At the other, there is direct manipulative interference through coercion, indoctrination, or neurological interference. Keeping people away from bad influences is at the mild end of the spectrum and is often unobjectionable. For example, restrictions on advertising of harmful tobacco products are designed to keep people away from bad influences for their sake, but this does not seem especially problematic.

Third, the extent to which shaping the life of another for their sake is problematic depends on the particular aspect of their life. Autonomy is not unconditionally valuable, or worthy of respect. There may be liberal restrictions on influencing people to develop some beliefs, tastes, or desires, or to value some things more than others. Within limits, it may even be wrong to influence people not to value things that are valueless, as some liberal neutralists believe. But there is no liberal objection to preventing a person from acting in a seriously wrongful way, because it is not valuable for a person to decide for herself to act in that way.<sup>11</sup> I will pick up this thread in Section VI.C, when we consider the relationship between the distribution of responsibility for wrongdoing and the distribution of other things.

The scope of distributive justice for responsibility, then, is controversial, where perfectionists and various different liberals will disagree about the legitimate aims that people may have in shaping the lives of others by shaping what they will be responsible for. But these disputes affect the scope and implications of the idea that responsibility is a matter of distributive justice, rather than the truth of that idea.

11. See, further, Victor Tadros, *Wrongs and Crimes* chap. 8.

## VI. Responsibility and Other Interests

Responsibility for wrongdoing is only one interest that makes social institutions or choices just or unjust. A complete account of the fair allocation and distribution of things that matter depends on the place of responsibility for wrongdoing among other justice relevant interests.

A. *Counterbalancing Responsibility*

Some think that welfare is the, or a, currency of distributive justice; others think it is opportunity for welfare; others capacities; and others resources. Still others think that the appropriate way to distribute resources is not to consider a metric for their distribution, but rather to engage in a hypothetical bargaining process that is appropriately sensitive to people's judgments and choices. These disputes are largely about what we owe to others with respect to, on the one hand, ensuring that they are well off, and, on the other, respecting their autonomy.

We can side-step these disputes. Suppose that some resources can be distributed between two people who will have equal capacities and opportunities to use any resources allocated to them. They have similar preferences and judgments, and they will use those resources to enhance their welfare. And, suppose, the magnitude and quality of their opportunities, as well as their actual welfare enhancement, is proportional to the magnitude of resources they are given. If other things are equal, on almost all views, the resources ought to be distributed equally between the two. If the resource available for allocation is indivisible, it ought to be allocated by some fair mechanism, such as a coin-flip on almost all views. I thus consider the interaction between the distribution of responsibility and the distribution of welfare-generating resources that are equally valued.

Now let us see what difference it makes to our judgments about how such resources should be distributed if we vary the distribution of responsibility for wrongdoing but hold other things equal. That will help us understand the way in which the distribution of responsibility for wrongdoing interacts with other matters of distributive justice.

Consider:

*Vision with Welfare:* As *Vision*. However, after having her vision, Teacher is provided with some indivisible resources, which she can either place on the seat next to Jack or the seat next to John. Whoever

sits on the seat with the resources on it will wrong Jeff. However, that person will use his extra resources at some point in his life to promote his welfare, and that the person will value this.

I have not specified the quality of the welfare that the person will gain. We could imagine a range of cases. Many people think that welfare is increased by (among other things) satisfaction at the exercise of virtues. So, let us suppose that each has an appreciation for art, but one will have greater resources and opportunities to visit galleries than the other, so that his welfare is increased. Let us then say that Teacher has an “art ticket” that she can place on either of the two seats. Should she place it on the seat next to Jack, the seat next to John, or flip a coin?

I find it clear that Teacher should place the ticket on the seat next to Jack, and a large majority of people that I have explored this issue with agree (although a small minority demur). This judgment is not affected by the fact that there may well be some determinism-compatible sense in which the person who is seated next to Jack could have refrained from acting wrongly. The fact that a person will act wrongly generates entitlements, even where the person could, in some sense, have refrained from acting wrongly.

Here is an explanation. Teacher, we have seen, ought to flip a coin to determine who sits next to Jack, where each child has an interest in not being seated there. We have seen that this interest is in not being responsible for wrongdoing. Thus, whoever is seated next to Jack is unlucky to be seated there. Teacher can counterbalance this bad luck by providing the unlucky person with something he has an interest in—the art ticket. In contrast, if Teacher places the art ticket on the seat next to John, whoever is seated next to Jack will be doubly unlucky in having two of his interests unsatisfied: his interest in not being responsible for wrongdoing and his interest in the art ticket. It is fairer that the full range of things that a person has an interest in are distributed more evenly than less evenly, and therefore Teacher ought to place the ticket on the seat next to Jack.

One way to characterize this view is that Teacher has a duty to counterbalance the bad luck of it being inevitable that a particular person is responsible for wrongdoing, if he is seated next to Jack, with the good luck of his being provided with welfare-generating resources. Or, to put it more controversially still, we ought to compensate people for being responsible

for wrongdoing by providing them with more welfare-generating resources than those who are not responsible for wrongdoing.

### *B. Intrapersonal Tradeoffs*

Some might object that although responsibility for wrongdoing and welfare are matters of distributive justice, they do not interact with each other. Those with this view might think that Teacher ought to flip a coin to determine who gets the art ticket in *Vision with Welfare*. This view is hard to believe. Why are we not morally required to consider all of the different interests that a person has in determining what we do for her, rather than treating these different interests completely independently?

This view that the different currencies of distributive justice interact is supported by the fact that in ordinary life, we regularly make intrapersonal tradeoffs between welfare-generating resources and responsibility for wrongdoing. Consider upbringing. Suppose that you have a child. Depending on how you bring her up, you will vary the amount of welfare-generating resources and responsibility for wrongdoing in her life. If you spend all of your time focusing on ensuring that she does not act wrongly, she will have fewer welfare-generating resources and will be less happy as a result. If you spend all of your time focusing on her welfare, she will act wrongly.

We commonly make decisions about how to bring up our children by balancing these considerations. Suppose, for example, that a small decrease in her responsibility for wrongdoing will result in her having far fewer welfare-generating resources: you can get her never to lie, but this will make her miserable throughout her life, where if she lies a little bit, she will be very happy, due to the resources she is provided with. You should sacrifice a bit of rightness for the sake of providing her with welfare-generating resources. Or, to take the converse case, suppose that a small decrease in her welfare-generating resources will result in her being far less responsible for wrongdoing. You should sacrifice some resources for the sake of her not being responsible for wrongdoing. This simple thought experiment shows that when we decide what to do for the sake of a person, we make intrapersonal tradeoffs between welfare and responsibility for wrongdoing, and simple moral decisions about these tradeoffs are uncontroversial.

The fact that we make judgments about what to do for a person by balancing their interests in not being responsible for wrongdoing with other

interests helps to support the view that when we consider a person's entitlements, we ought to consider their overall set of interests together. And when we compare their entitlements with those of other people, we compare their overall sets of justice relevant interests with each other. We counterbalance the failure to satisfy a person's interests in welfare-generating resources by satisfying her interests in not being responsible for wrongdoing and vice versa. Normally, where there is a competition between the interests of two people, and we disadvantage a person when compared with another with respect to one kind of interest, we have reason to advantage her with respect to another.

### *C. Responsibility and Judgment*

Another objection to the idea that people should be given welfare-generating resources to counterbalance responsibility for wrongdoing draws on the idea that we are not required to compensate people for things they are responsible for, or that respond to their authentic judgments, which is a familiar theme in recent debates on distributive justice.

Here is one version of the objection. Suppose that one person, A, has tastes that are much more expensive to satisfy than another, B. It might be argued that distributive justice does not require us to distribute resources in a way that equalizes the welfare of A and B, other things equal. It is unobjectionable that B is left better off than A where that difference is due to A's expensive tastes. The taste for wrongdoing, it might be argued, is like an expensive taste, in that it inflicts costs on others.

This issue merits a longer discussion than I have space for here, but I can at least sketch a response. First, the view that we should not compensate people who have less welfare as a result of their expensive tastes is itself a controversial view. To determine whether compensation is owed, we should, among other things, consider whether the initial judgment that such tastes should not be compensated can survive critical scrutiny in the light of the kinds of deterministic challenges outlined in this article. Some think that the question whether responsibility-sensitive egalitarianism is right depends on whether there is genuine responsibility.<sup>12</sup> This article suggests a different challenge—that even if there is genuine responsibility,

12. For anxiety about whether the significance of choice in distributive justice can survive such scrutiny, see, for example, G. A. Cohen, *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy* (Princeton: Princeton University Press, 2011), pp. 32 and 119–20.

a brightly illuminated understanding of what responsibility is, and its social causes, gives it a quite different role in distributive justice than responsibility-sensitive egalitarians suggest.

To make progress, responsibility-sensitive egalitarians might consider versions of the *Vision* cases where different people have more or less expensive tastes depending on where they were seated in childhood. Investigating these cases is beyond the scope of this article, but it is by no means obvious that the initial judgment that we should not compensate people for expensive tastes will survive scrutiny in the light of such an investigation.

Furthermore, at least some of the argument that has been given against compensating people for expensive tastes does not apply to responsibility for wrongdoing. Very roughly, there are two kinds of argument that might be given. One is that people are not owed compensation because a liberal society distributes resources without making judgments about the good life, out of respect for its citizens. The second is that people are not owed compensation for expensive tastes because if they are worse off, this will be their own fault, or because they had an adequate opportunity to avoid being worse off.

The first argument does not apply to responsibility for wrongdoing. Consider Ronald Dworkin's influential approach to the distribution of resources.<sup>13</sup> Dworkin argues that liberal societies are built on equal concern and respect. Respect for citizens requires such societies not to distribute resources based on an evaluation of how well a person's life goes based on any particular conception of the good. Rather, the appropriate way to determine what resources a person is entitled to is to consider whether a person envies the bundle of resources that others have. In that way, such a society has a fair mechanism for determining who gets what whilst refraining from making judgments about the quality of a person's own life.

The person with expensive tastes need not envy the resources and opportunities of those with less expensive tastes. She will not do so if she values her expensive tastes over less expensive tastes. Even if she is left off worse in welfare terms than the person with less expensive tastes, she will not find the resources that she is provided with objectionable compared

13. See Ronald Dworkin, *Sovereign Virtue* (Cambridge, Mass.: Harvard UP, 2000); Ronald Dworkin, *Justice for Hedgehogs* (Cambridge, Mass.: Harvard UP, 2011), chap. 16.



with the resources provided to others. Dworkin proposes a particular set of mechanisms—a hypothetical auction coupled with a hypothetical insurance scheme—to make this general idea concrete.<sup>14</sup> The details of that scheme need not detain us here.

This view does not support the view that people should not be compensated responsibility for wrongdoing. Here are two replies that draw on Dworkin's ideas—one more limited and one more ambitious. The more limited reply is that the distinction between facts that ground a claim for resources and facts that do not depends on what people authentically endorse about their lives according to their own values. Dworkin distinguishes between expensive tastes that reflect a person's values and mere cravings and disabilities that the person does not endorse.<sup>15</sup>

This can create the misleading impression that if a person is responsible for being worse off, she is not entitled to compensation. But what a person is responsible for and what a person authentically endorses in the light of what she values are not coextensive. A person can be responsible for her conduct without endorsing or valuing it, and many people who are responsible for what they do fail to value what they do. A person who gives into temptation, and acts wrongly as a result, for example, does not value acting wrongly, and respect for her does not militate against compensating her for acting wrongly. Furthermore, a person may value acting wrongly, but do so because she fails to draw proper conclusions from what she values, due to temptation, manipulation, or social pressure. Respect does not require us to withhold compensation where a person fails to conclude that she is worse off for these reasons, but they need not undermine responsibility.<sup>16</sup>

This reply, though, leaves open the possibility that a person would not be entitled to compensation for responsibility for wrongdoing in cases where she authentically values acting wrongly. But it is doubtful that Dworkin's view, as it is best understood, denies the right to compensation even in these cases.

Consider the idea that in a liberal society, we defer to a person's own judgments about what is good for her out of respect for her as an

14. See, especially, Dworkin, *Sovereign Virtue*, chap. 2.

15. *Ibid.*, pp. 287–96; Ronald Dworkin, "Ronald Dworkin Replies," in *Dworkin and His Critics*, ed. J. Burnley (Oxford: Blackwell, 2004), pp. 346–50.

16. See, also, Tom Parr, "How to Identify Disadvantage: Taking the Envy Test Seriously," *Political Studies* 66 (2018): 306.

autonomous agent. If this is the foundation of a person's claim to resources, we should explore when respect requires us to defer to the judgments of others. This will not provide a wholesale justification for deferring to the views of citizens about what to value. And that makes a difference to when compensation is owed.<sup>17</sup>

One obvious limit to the idea that we should respect a person's judgment with respect to how they live their own life concerns responsibility for wrongdoing. Respect for a person does not require us to defer to her judgment that serious wrongdoing is valuable for her. Think about the idea that I might have reason to value having lived my own life according to my own values. This is true when the what I value is valuable. Perhaps it is also true where what it is reasonable to value what I value. But it is surely not true where what I value is wronging others. Then, I have reason to disvalue having lived my life according to my own values. That helps to vindicate the intuition that where I find out that I have acted wrongly, I have reason to hope that I did not do so autonomously. Thus, respect for agents in virtue of what they value does not provide a reason to refrain from providing wrongdoers with compensatory resources for acting wrongly, even where they value acting wrongly.

To justify a policy of refraining from compensating those who are responsible for wrongdoing, then, we cannot appeal to the idea that the distribution of resources in a society is governed by respect for autonomous agents, and the different things they value in their own lives. It must be justified by something like the second kind of argument: that a person is not entitled to compensation if their being worse off is their own fault, or if they had an adequate opportunity to avoid acting wrongly. And that is precisely the kind of view that the *Vision* cases challenge. When we clarify the social and natural causes of responsible conduct, we see fault as a reason to transfer resources to the wrongdoer, not away from her. And the fact that she had an opportunity to avoid doing what she did does not undermine that judgment where what she did was inevitable given the state of the world, and the laws that apply to it, before she was a responsible agent.

17. See, especially, Adam Slavny, "On Being Wronged and Being Wrong," *Philosophy, Politics & Economics* 16 (2017): 3.

#### D. Extra Options

Some may think that the reason why teacher ought to put the art ticket on the seat next to Jack in *Vision with Welfare* is that the fates of Billy and Bobby are codependent. Billy being placed next to Jack makes it the case that Bobby will not be responsible for wrongdoing and vice versa. Lack of responsibility for wrongdoing is like a scarce resource in this case in the sense that the satisfaction of one person's interest ensures that another person's interest is not satisfied. It might be argued that welfare-generating resources ought not to be distributed to those who are responsible for wrongdoing where this is not so—where one person's responsibility for wrongdoing does not depend on another's. This, it might be thought, limits the significance of *Vision with Welfare* in the real world, where lack of responsibility for wrongdoing is not a scarce resource.

One response is that in the real world, lack of responsibility for wrongdoing is often a scarce resource. Preventing responsibility for wrongdoing takes investment in the kinds of programs that we considered in Section II, but the investment needed for these programs is scarce, and we must decide where to allocate these preventive resources.

More importantly, the proper allocation of welfare-generating resources in *Vision with Welfare* is not best explained by the fact that the fates of Billy and Bobby are codependent. Consider:

*Assistant's Vision with Welfare and Other Options:* As Seating Arrangement except Teacher has the option of sitting neither child next to Jack—there are two empty seats either side of John, and if the children are seated in these seats, neither will act wrongly later in life. However, Assistant has a vision of the future. She knows that one person will be seated next to Jack and the other will be seated next to John. Furthermore, the second person, whoever that is, would be seated next to John even if the first were not seated next to Jack—there is another empty seat next to John that the first would have taken. Assistant is provided with some indivisible resources, which she can either place on the seat next to Jack or the seat next to John.

Teacher makes the objectively wrong choice in seating a child next to Jack—she should seat both children next to John. But the first child being seated next to Jack does not prevent the second from being responsible

for wrongdoing. Yet it is still powerfully intuitive that Assistant should place the welfare on the seat next to Jack. Thus, welfare-generating resources should be allocated to those who are responsible for wrongdoing even where lack of responsibility for wrongdoing is not a scarce resource. This is true whether Teacher is culpable for her choice to seat a child next to Jack, because she has good reason to predict what will happen, or not.

And this case has a further important implication. Assistant has a powerful reason to counterbalance responsibility for wrongdoing even though Teacher's conduct, not hers, results in a person being responsible for wrongdoing. Thus, third parties should counterbalance responsibility for wrongdoing that is caused by others with welfare-generating resources.

### *E. Time*

Another response to *Vision with Welfare* is to accept that the art ticket ought to be placed on the seat next to Jack, but to deny that this has implications for how a person should be treated later in life. Some might claim, for example, that we have reason to advantage those who will inevitably be responsible for wrongdoing prior their acting wrongly, but not after they have acted wrongly.

There are two time-sensitive views. An agent-centered view is that we are required to counterbalance responsibility for wrongdoing with welfare-generating resources only where we can act prior to the person being responsible for wrongdoing. A patient-centered view is that we are required to counterbalance responsibility for wrongdoing only where the allocation of resources, or the effects of those being allocated, will occur before the person is responsible for wrongdoing.

Neither view is attractive. Focus on the agent-centered view first. Suppose that I can provide a wrongdoer with more welfare-generating resources after she has acted wrongly. On the agent-centered view, I am required to do this if my act that causes her to acquire or use those resources occurs prior to her wrongdoing, but I am required not to do it if my act that causes her to acquire or use them occurs after her wrongdoing. It is hard to see why this should make a difference.

Some might argue that many of our responses to wrongdoing are inherently backward looking, and that this might support the agent-centered view. For example, it is not apt for me to resent you, or to demand an

apology and compensation, until you have wronged me, even where I have powerful reasons to believe that you will wrong me. These reactions to wrongdoing are irreducibly backward looking.

However, even if there is an important relationship between my reactions and demands and my distributive decisions, the main reason why these attitudes are backward looking is epistemic. The aptness of my having the relevant attitudes and making the relevant demands depends on my knowing about your wrongdoing, and I cannot normally know that you will act wrongly.

I can sometimes know that you will act wrongly, but then it does seem apt to have the relevant attitudes and make the relevant demands. For example, suppose that you have a clear plan to wrong me, you are about to wrong me, nothing stands in your way, and you have a strong track record of following through. Might I not then know that you will wrong me? I think I might. But if I do, it is apt to resent you for the wrong that you will do. Suppose, for example, that you are in the course of beating me to death. It seems apt to resent you for the fact that you will kill me. And it seems apt to make demands of you in virtue of the fact that you will—for example, the demand to help my family.

This demand might seem odd—it's more important to demand that you don't kill me. But even if it is odd in this way, making the demand that you respond to your wrongdoing is not inapt because it is made too early. And it need not be odd—if I know that I cannot dissuade you from acting wrongly, I might have good reason to demand that you respond appropriately to your wrongdoing rather than demanding that you desist. As the limits on the aptness forward-looking responses to wrongdoing are primarily epistemic, we cannot rely on them to support the agent-centered view; if anything, they support the time-neutral view that it does not matter when we make our distributive decisions.

The patient-centered view fares no better. On that view, whether Teacher should put the art ticket on the seat next to Jack or the seat next to John depends on whether the ticket is valid prior to or after the wrongful conduct. That is unintuitive in itself. Furthermore, the tradeoffs we make between responsibility for wrongdoing and welfare-generating resources in the case of upbringing are not time-sensitive. When deciding the appropriate balance between welfare-generating resources and responsibility for wrongdoing in upbringing, we do not consider the order

in which these things will occur. It is hard to see why we should draw a different conclusion in interpersonal cases.

Here is my conclusion. If I am right that if Billy is seated next to Jack, Teacher ought to place the art ticket on the seat next to Jack, she ought to give Billy the art ticket later in life in World 1, whether her act of giving, or Billy's receipt of the ticket, occurs prior to or after he wrongly assaults Jeff. So now suppose that she is in World 1, sees Billy assault Jeff, and knows that her decision to sit him next to Jack made it true that he assaults Jeff. If she has an art ticket that she can give to either Billy or Bobby, distributive justice favors giving it to Billy other things equal. Furthermore, we can conclude from this discussion, and our judgment about *Assistants Vision with Welfare and Other Options*, that if third parties know that Teacher's decision has this effect, distributive justice also favors their giving it to Billy other things equal.

Now reflect on the fact that all responsibility for wrongdoing is shaped by social decisions such as those in the variations on *Seating Arrangement* that we have been considering, and we have significant support for the idea that distributive justice favors providing extra welfare-generating resources to everyone who is responsible for wrongdoing in the real world. And it is a short step to the view that distributive justice favors providing fewer welfare-generating resources to everyone who is responsible for good deeds in the real world.

#### VII. Natural and Social Causes

Some may reject the conclusion drawn in the previous section out of a further concern. I have argued that distributive justice requires us to distribute welfare-generating resources to those who are responsible for wrongdoing, and away from those who are not, where that wrongdoing arises as a result of social decisions, other things equal. It seems that all wrongdoing is attributable to social decisions, so this view seems to require us to distribute welfare-generating resources to everyone who is responsible for wrongdoing, and away from those who are not, other things equal. However, it might be thought that things are different where responsibility for wrongdoing arises through natural differences between people, and they often do in the real world.

Those inclined to this view might draw comfort from the fact that some think that inequalities more generally are unjust where they arise socially,

but not naturally.<sup>18</sup> Suppose, for example, there are no social differences between the two seats in a variation on *Vision with Welfare*, but Billy and Bobby act differently due to genetic differences between them. Then, some might conclude, there is no reason to give Billy rather than Bobby extra welfare-generating resources.

I doubt that the distinction between natural and social causes of responsibility for wrongdoing is morally significant. Familiar and powerful challenges to the significance of that divide extend to the just distribution of responsibility for wrongdoing. One question concerns the best way of drawing the social/natural distinction. In the real world, all responsibility for wrongdoing arises due to some social facts. There are no purely natural sources of inequality between people with respect to responsibility for wrongdoing. So, our question is whether there is a way of drawing the natural/social distinction in a way that is plausible, and that has implication in our context. In the context of distributive justice more generally, Thomas Nagel suggests that we do this as follows: where institutional arrangements serve a purpose that is not to generate inequalities, and inequalities arise due to natural differences, and achieving the purpose without generating inequalities would be more difficult and costly, the inequalities are natural, and are thus not unjust.<sup>19</sup>

But this view is not plausible.<sup>20</sup> To see the problem clearly in our context, compare *Vision (Tone Variation)* with a case where Jack and John have identical voice tones, but Billy and Bobby are genetically different, so that Billy is more likely to respond to their voice tone by doing what he is told than Bobby (*Small Genetic Variation*). There does not seem to be any important difference between *Tone Variation* and *Small Genetic Variation*. And there are powerful reasons to ensure that responsibility for wrongdoing is distributed equally between Billy and Bobby to eliminate the effects of genetic differences between them because of their tendency to respond to these different voice tones, even where doing this is costly.

18. See, especially, Thomas Nagel, "Justice and Nature," in *Concealment and Exposure and Other Essays* (Oxford: OUP, 2002). A variation on this view is that justice is only concerned with inequalities that have an institutional cause. See Kok-Chor Tan, *Justice, Institutions, and Luck* (Oxford: OUP, 2012).

19. Nagel, *Concealment and Exposure and Other Essays*, p. 125.

20. For a range of concerns in more familiar distributive contexts, see Tim Lewens, "What Are 'Natural Inequalities'?", *Philosophical Quarterly* 60 (2010): 264; Kasper Lippert-Rasmussen, "Are Some Inequalities More Unequal than Others? Nature, Nurture and Equality," *Utilitas* 16 (2004): 193–219.

Perhaps it might be argued that things are different in cases where one person is naturally much more robustly disposed to be responsible for wrongdoing than another. Suppose, for example, that Billy is genetically disposed to be responsible for wrongdoing, so that he will act wrongly in a much wider range of nearby possible worlds than Bobby, say because Billy has higher levels of testosterone than Bobby (*Large Genetic Variation*).

It is hard to see why this makes a difference. Perhaps it might be argued that there are stronger reasons for Billy to bear costs where these costs arise as a result of features of him that stably incline him to act wrongly. But it is hard to see why this is a reason to make him worse off than others, who are caused to be responsible for wrongdoing in the actual world, but whose tendency to act wrongly is less stable. Perhaps it might be argued that this is true where larger genetic differences between people also lead them to develop different characters, so that they endorse their conduct. But this just brings us back to an issue we considered earlier: whether respect for a person's values makes a difference to what we should do for them when they are morally worse off. I suggested that even if it sometimes does, this has no bearing in the case of responsibility for wrongdoing.

The same conclusions hold when we consider the implications of larger social decisions that affect responsibility for wrongdoing. Suppose that there are different ways to plan cities. One way of doing this (Plan 1) will result in smaller inequalities in responsibility for wrongdoing, where the other (Plan 2) will result in larger inequalities of this kind. However, Plan 2 is more costly, so everyone's welfare-generating resources are slightly reduced if Plan 2 is implemented. Whether Plan 2 should be implemented does not depend on whether the resulting inequalities arise due to the interaction of genetic facts and features of the plan, or whether they are purely social. I conclude that distributive justice supports distributing welfare-generating resources to everyone who is responsible for wrongdoing, other things equal, regardless of the social or natural origins of their responsibility for wrongdoing.

#### VIII. Why Distributive Justice Is not the Whole Story

I have argued that distributive justice requires distributing welfare-generating resources to those who are responsible for wrongdoing, other things equal, and away from those who are responsible for do-gooding. Of



course, other things are often unequal. Distributing welfare-generating resources to wrongdoers incentivizes wrongdoing, and that is a powerful reason not to do this. And one reason is that incentivizing wrongdoing gives rise to the very inequalities that we have been concerned with—by attempting to correct the inequality between one person who is responsible for wrongdoing and another who is not, we may cause a third person to act wrongly, who will then be a victim of the very kind of injustice that we aimed to correct.

Furthermore, distributive justice is not the only consideration that determines how resources ought to be allocated. It is part of the overall set of factors that we take into consideration when deciding what to do. A full assessment of the relationship between distributive justice and other kinds of duties and permissions is beyond the scope of this article, but I can at least illustrate the general idea that there are limits to the significance of distributive justice for our duties with respect to wrongdoers.

Here is one idea, though it is more limited than it at first seems. Justice requires that we adequately respond to wrongdoing by vindicating the rights of victims and condemning wrongdoing through public expression. But, it might be argued, public expression would be adequate only if we inflict costs on wrongdoers.<sup>21</sup> And it would be undermined were we to provide welfare-generating resources to wrongdoers.

Quite generally, I doubt that inflicting costs on wrongdoers is necessary for the relevant forms of public expression.<sup>22</sup> The aim of expressing condemnation of wrongdoing, and vindicating rights of victims, only seems undermined if welfare-generating resources are provided to those who are responsible for wrongdoing in certain ways. Just distributing cash to those who are responsible for wrongdoing, for example, might be difficult to achieve without creating the impression that the state is unconcerned with the wrongdoing, or even values it.

21. This idea is familiar from expressivist theories of punishment. See, for example, Jean Hampton, "The Retributive Idea," in *Forgiveness and Mercy*, eds. J. Murphy and J. Hampton (Cambridge: CUP, 1988); Joshua Glasgow, "The Expressivist Theory of Punishment Defended," *Law and Philosophy* 34 (2015): 601; Bill Wringe, *An Expressive Theory of Punishment* (Basingstoke: Palgrave, 2016).

22. See Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford: OUP, 2011), chap. 5.

However, suppose that a state significantly increased funding of rehabilitative and social integration programs, and these were very effective both in ensuring that wrongdoers recognized that their conduct was wrong, and ensured that those responsible for serious wrongs engaged in productive and rewarding work later in life. That proposal provides those who are responsible for wrongdoing with more welfare-generating resources but does not undermine the legitimate expressive or communicative ambitions of the state. Still, there may be limits, in some social conditions, to what can be done to promote distributive justice to those who are responsible for wrongdoing, without undermining these ambitions.

Here is another idea. The permissibility of inflicting a cost on a person for the sake of some goal depends on the duty that the person has to pursue the goal at the relevant cost. This makes a difference to the costs that can be inflicted on a person for the sake of assisting those they wrongly harm. A person who acts wrongly may incur the relevant duty, and that may make it permissible to inflict the cost on her, even where inflicting this cost results in a distributive unfairness.

Consider, for example, whether Billy or Bobby owes Jeff assistance after Billy assaults him in World 1 in *Seating Arrangement*. I have argued that distributive fairness militates in favor of allocating resources to Billy rather than Bobby. But that may not fully settle the question of who is required to assist Jeff. Billy's special connection to the harm that Jeff suffers because he is responsible for it may give rise to a more stringent obligation on him to assist Jeff.<sup>23</sup> His having this duty gives rise to a further distributive unfairness: he will both be responsible for wrongdoing, and, if he does what he is required to do, he will be worse off than Billy with respect to welfare. But it does not follow that Bobby has a more stringent duty to assist Jeff than Billy. Distributive fairness and unfairness does not fully determine our duties.

To see that this is so, consider the less controversial idea that people have more stringent duties than strangers to help their family members. The fact that my parent requires a great deal of assistance, where other people's parents do not, may result in my being worse off than others. Some may believe that this gives rise to duties on others to compensate

23. For an argument that such duties survive in a deterministic world, see Tadros, *Wrongs and Crimes*, chap. 5. Some prominent free-will skeptics accept this. See Derk Pereboom, "Self-Defense, Deterrence, and the Use Objection: A Comment on Tadros's *Wrongs and Crimes*," *Criminal Law and Philosophy* 13 (2019): 439.

me to ensure distributive justice. I think that there are limits to this duty, though, and sometimes my more stringent duty to my parents will leave me worse off than others. Where this is so, others lack a duty to achieve distributive justice all things considered.

The same is true of Billy and Bobby. Billy owes a stringent duty to assist Jeff for having assaulted him. Fulfilling this duty will leave him worse off than Bobby. Because this is distributively unfair, Bobby has some duty to assist Billy. But this duty may be insufficiently stringent to achieve distributive fairness between them. In other words, two different considerations pull in different directions. Billy's special connection to the harm that Jeff suffers militates in favor of his having a more stringent duty to assist Jeff, even where that will leave him worse off. But distributive fairness militates in favor of Bobby having a duty to bear costs to assist Jeff. This may result in a situation which leaves Billy worse off than Bobby overall.

#### IX. Incompatibilism to the Rescue?

I have argued from compatibilist assumptions for a radical view in distributive justice—that distributive justice favors allocating welfare-generating resources to those who are responsible for wrongdoing, and away from those who are responsible for do-gooding. Although I softened the blow of this radical conclusion somewhat in Section VIII, where I suggested that distributive justice is not the whole story, many will find my conclusion unpalatable. Where the conclusion is more counterintuitive than the premises, some might argue, the right response is to reject one of the premises. I will raise some doubts about this idea in Section X.

There is one important premise that I haven't argued for: compatibilism. I just made some (fairly modest, it should be said) compatibilist assumptions in Section I. Although compatibilism is the most popular view of responsibility, it has plenty of critics. Although I cannot argue for compatibilism here, though, I can give reasons to think that my argument, or a similar argument, will be compelling on any plausible incompatibilist view.

Here's why. Incompatibilists believe that responsibility is incompatible with determinism. Now suppose that *Determinism* is false. Some incompatibilists believe that the fact that it was not inevitable that I acted in a certain way in the actual world might make me responsible for my conduct. However, any plausible version of this view must make room for

the fact that people's conduct is subject to social influences, and these influences increase or decrease the probability of wrongful or do-gooding conduct. To be plausible, incompatibilism must be able to explain, for example, why advice and incentives increase the probability of people acting one way or another without undermining her responsibility for her conduct. For surely a person can be responsible for what she does even though she acts as a result of advice and incentives that increase the probability of the conduct she is advised or incentivized to perform.

But social facts alter the probability of all of our conduct. People then have an interest in the existence of social facts that decrease the probability of their acting wrongly and increase the probability of do-gooding. This has distributive implications that are similar to those explored earlier. For example, suppose that in a variation on *Vision (Three Options)*, Teacher knows that a child that sits next to Jack has the highest probability of being responsible for wrongdoing, a child that sits next to John has the lowest probability of being responsible for wrongdoing, and a child that sits next to Billy or Bobby has an intermediate probability of being responsible for wrongdoing. Distributive justice favors the third option of sitting Billy and Bobby together.

It might be argued that this view does not establish that responsibility for wrongdoing is part of the currency of distributive justice; only that facts that affect the probability of responsibility for wrongdoing are part of the currency of distributive justice. But although justice requires the fair distribution of the probability of responsibility for wrongdoing, that element of justice is completely dependent on the significance of responsibility for wrongdoing. And the outcome with respect to responsibility for wrongdoing ultimately determines how future decisions about welfare-generating resources should be made.

Suppose that Teacher sits Billy next to Jack. This is against Billy's interests, because it increases the objective probability that he will be responsible for wrongdoing. If Teacher has an art ticket, she ought to give that ticket to Billy, some might think. That is true where Teacher does not know the outcome of her decision. Then, by allocating the ticket to Billy, she increases the probability that it will go to the person who is responsible for wrongdoing. But now suppose that the ticket is to be allocated after the time where he might act wrongly, but he does not. Now Teacher lacks a reason to allocate the art ticket to Billy. Why? Because his interest in not having his chance of being responsible for wrongly increased is completely

derived from his interest in not being responsible for wrongdoing. Where the increased chance of his acting in this way does not actually result in responsibility for wrongdoing, there is no reason to counterbalance the unfair distribution of chances by providing him with welfare-generating resources.

Furthermore, suppose that although Billy has a greater chance of being responsible for wrongdoing than Bobby because of where they are seated, Bobby acted wrongly. Then distributive justice militates in favor of providing welfare-generating resources to Bobby. Although the unequal distribution of chances favored him *ex ante*, this did not in fact benefit him. The result of the actual role of the dice determines what we should do, even where the dice are weighted.

To reinforce this view, compare the relationship between risks and outcomes in other distributive contexts. Suppose that I have an indeterministic machine that will allocate welfare to Billy and Bobby. The machine has only one button, which gives Billy a 75% chance of receiving a benefit and Bobby a 25% chance of receiving a benefit. If I have a further welfare-generating resource that must be allocated prior to pressing, I ought to give it to Billy. Now suppose that I can only allocate the resource after pressing. I press the button, and the machine gives the benefit to Bobby. I now have reason to benefit Billy, and no significant reason to benefit Bobby; the fact that Billy was better off with respect to the chance of receiving a benefit seems to make little or no difference when deciding how to act in the future.

A similar thing is true with respect to responsibility for wrongdoing. Suppose, for example, that Billy and Bobby are each torn, where that the physical feature that makes it true that they act wrongly, or not, is an indeterministic mechanism in their brains. But the mechanism gives Billy a 25% chance of acting wrongly where Bobby has a 75% chance of acting wrongly. However, the mechanism results in Billy acting wrongly where Bobby does not. I should allocate welfare-generating resources on the basis of the outcome, not the probabilities, if possible. In an indeterministic world, we have reason to distribute resources or welfare to those who subject to the social or natural conditions that make them more likely to be responsible for wrongdoing *ex ante*. But we only have reason to distribute resources or welfare to those who are actually responsible for wrongdoing *ex post*.

This, of course, is only a sketch of an argument. But it at least suggests that the strongly counterintuitive view that distributive justice requires us

to distribute welfare-generating resources to those who are responsible for wrongdoing is true if standard folk judgments about responsibility are true, regardless of whether they are best explained by compatibilist or incompatibilist theories of responsibility.

#### X. Conclusion

Some may still be unmoved. The idea that distributive justice favors providing welfare-generating resources to wrongdoers, and away from do-gooders, is, I have acknowledged, very strongly counterintuitive. Some might think that even if I have provided significant support for the premises of my argument, rejecting them is less counterintuitive than accepting the conclusion.

Here is a response. When considering the extent to which an intuition supports a view, we should consider both its strength and its reliability. This is familiar from the analysis of manipulation cases in debates about responsibility. A powerful strategy for compatibilism skeptics is to begin with cases where a person is manipulated by evil scientists to act wrongly, but fulfills all plausible compatibilist conditions of responsibility. The skeptic then aims to show that there is no responsibility-relevant difference between manipulation cases and everyday cases where a person's conduct is fully determined without manipulation.<sup>24</sup> A familiar response is that this test can be run backward—start with cases not involving manipulation where it is strongly intuitive that we are responsible, show there is no difference between these cases and manipulation cases, and draw the conclusion that we are responsible in manipulation cases.<sup>25</sup>

This seems to create a standoff between skeptics and compatibilists. But at this point each can make claims about the strength and the reliability of intuitions about these cases. On the one hand, following Benedict de Spinoza, skeptics point out that we have unreliable intuitions about free will that result from the fact that, from the internal point of view, our conduct does not seem fully determined, and the causes of our conduct are not fully present in our decisions.<sup>26</sup> On the other hand, compatibilists point out that we have unreliable intuitions about manipulation cases,

24. The best argument of this kind that I know of is Pereboom, *Free Will, Agency, and Meaning in Life*, chap. 4.

25. See, especially, McKenna, "A Hard-Line Reply to Pereboom's Four-Case Argument."

26. See, especially, Pereboom, *Free Will, Agency, and Meaning in Life*, pp. 92–93.

because they create the impression of a contrast between the fully responsible manipulators, and the person manipulated.<sup>27</sup> Thus it is difficult to know who wins.

There is no similar standoff with respect to our cases. Although the intuition that justice does not require giving welfare-generating resources to wrongdoers is powerful, it is also unreliable for Spinozist reasons, and perhaps for other reasons—some might think, for example, that our retributivist sentiments are vulnerable to evolutionary debunking arguments. But there is less reason to doubt the reliability of our intuitions about the *Seating Arrangement* cases. The compatibilist response to manipulation cases is not available to challenge those intuitions, and it is not easy to see what other doubt we should have. Perhaps it might be argued that we generally find it hard to hold in our head both the idea that a person's conduct is determined by a background choice, such as Teacher's, and the (compatibilist friendly) voluntary choice that the person makes. I doubt that this is especially difficult—it is a common feature of everyday life that our actions can be shaped by external factors, such as persuasion and advice, and yet chosen.

Even if my conclusion is more strongly counterintuitive than some of the judgments about cases that I rejected in my arguments, intuitions that conflict with the conclusion are less reliable. Overall, I am more confident in the radical conclusion than I am in the view, for example, that the art ticket should be placed on the seat next to John in *Vision with Welfare*. We may be responsible for wrongdoing, but that is a reason to make us better off.

27. See Sartorio, *Causation and Free Will*, pp. 156–70; Tadros, *Wrongs and Crimes*, pp. 76–77.