



Moral Luck, Responsibility, and Systems of Tort Liability

Emmanuel Voyiakis¹

© The Author(s) 2020

Abstract

Bernard Williams drew our attention to what might be wrong with denying the role of luck in our understanding of agency and responsibility. Susan Wolf and David Enoch, in separate works, have asked us to focus instead on what might be virtuous and valuable in embracing that role, and on how our institutions might assist us in that regard. They claim that the agent who ‘takes’ a responsibility that law or morality do not already assign to them may be displaying a special moral virtue or exercising a distinctive moral power. I raise some objections to Wolf’s and Enoch’s case for that claim, and query some of its purported institutional implications for tort law systems.

Keywords Luck · Responsibility · Tort law · Liability · Bernard Williams · Agency · Morality · Institutions

Introduction

Debates about moral luck have gripped the interest of tort theory because they promise insights into the conditions under which our institutions may require an agent to make repair for accidents or other harmful events.¹ It seems beyond dispute, and I will assume that it is, that our institutions may do this only when the accident is connected to that person’s agency in the appropriate way. Arguments about moral luck matter for tort law precisely because they help us think about what counts as an appropriate connection between an action (or event) and an agent in relation to questions of responsibility for repair. My aim in this paper is to trace and discuss a particular line of thought about luck and agency, and the significance of their relationship between the two for the justification of systems of tort liability.

¹ From a large literature, see Honoré (1988), Waldron (1995), Perry (2001), Keating (2006), Goldberg and Zipursky (2007), Jackson (2012), and Gardner (2017).

✉ Emmanuel Voyiakis
e.voyiakis@lse.ac.uk

¹ Law Department, London School of Economics and Political Science, London, UK

My guiding illustration will be Bernard Williams's simple and powerful example of the lorry driver who, through no fault of his own, runs over a little child. Williams notes that although we feel sorry for the driver, 'that sentiment co-exists with, indeed presupposes, that there is something special about his relation to the happening, something which cannot merely be eliminated by the consideration that it was not his fault'.² It is the persistence of this connection that, for Williams, reflects the contribution of luck to the driver's moral profile.

Susan Wolf has elaborated on this intuition by inviting us to consider our reactions in two variations of Williams's example.³ Two drivers culpably increase the risk to the child by failing to check the brakes of their respective lorries before setting off. Neither of them could have stopped in time, even if their brakes had been functioning well. One driver hits the child, feels immense guilt, steps up and declares himself responsible for the accident. The other driver avoids the child, thanks his lucky stars for the close escape, and otherwise treats the event as having nothing to do with him, morally speaking. Wolf argues that, while neither driver could have avoided the accident, we will judge them very differently. In particular, she claims that what sets the first driver apart is that he displays a special moral virtue in *taking* responsibility for an accident that arguably lies outside the 'core' of his agency, understood as the sphere of actions and events over which he has a measure of control.⁴

David Enoch has developed Williams's intuition in a similar direction. Like Wolf, he sees claims about moral luck as claims about bringing an action or event from the penumbra of one's agency into the core, though he sees such 'taking' of responsibility as instantiating not a moral virtue but a moral power (and, occasionally, a moral duty).⁵ Even more helpfully for my present purposes, Enoch has worked through certain implications of that claim for tort liability. He argues that a tort system which failed to facilitate the exercise of the moral power of taking responsibility would be morally deficient for that reason.⁶ In his view, that is what might be missing from systems that meet the cost of accidents through public insurance schemes, like the New Zealand Accident Compensation Corporation, rather than through more traditional private law remedies.

My aim in this paper is to try to understand those claims about moral luck better, and to express certain doubts about their institutional implications. Besides having breathed new life into an old and somewhat tired debate about moral luck, the idea of 'taking responsibility' strikes me as worthy of close attention for two reasons. First, it treats moral luck not as a sceptical construct or puzzle for us to reason our way out of, but as a positive contribution to our moral imagination about agency and responsibility. Second, the implications of Enoch's claims for institutional action seem to me to add bite to certain aspects of Williams's own thoughts about the

² Williams (1981) at 28.

³ Wolf (2013).

⁴ Ibid. at 13–14.

⁵ Enoch (2012).

⁶ Enoch (2014).

significance of the availability of insurance coverage for accidental harms, and the connection between public systems of insurance and individual moral agency.

My paper makes two claims. The first, which echoes arguments made against more familiar interpretations of the idea of moral luck, is that our reasons for requiring the driver to bear certain responsibilities, both in Williams's original example and in Wolf's variations, can often be accounted for on epistemic grounds, without the need to hypothesise that those reasons track a moral virtue or a special moral power. For example, we have sufficient reason to treat the driver as responsible for the accident insofar as their response to events obstructs us from establishing what happened, or aims to put them in the clear too easily. My second claim is that, epistemic concerns aside, seeing the 'taking' of responsibility as the exercise of a moral virtue or power, and calling on our institutions to accommodate that exercise, should not lead us to assume that our ideas about agency have a certain priority over our ideas about institutional action and design. Perhaps the New Zealand system, which protects people from personal liability for certain accidental harms, is nothing short a laboratory of good moral ideas about what people should regard themselves and others as responsible for. Just like a good safety feature makes it unnecessary for people to go over a certain point of effort in keeping themselves and others safe, perhaps the New Zealand system makes it unnecessary for people to extend their agency in the ways that Wolf and Enoch envisage. Generally stated, my suggestion will be that the best account of our agency should build the solutions that institutions can provide into our moral ideas about who we are and what we are responsible for, rather than treat those solutions as mechanisms for implementing conclusions reached in a pre-institutional moral calculus.

I proceed in two steps. 'Williams' section summarises Williams's thoughts on the connection between accidental harms and conceptions of agency, and finds the aspects of those thoughts that relate to institutions and their proper set-up rather modest and guarded. 'Wolf and Enoch' section discusses Wolf's and Enoch's accounts, and asks how far we need to hypothesise a moral virtue or power of 'taking responsibility' in order to justify the moral reactions that Wolf and Enoch regard as appropriate in the variations of the lorry driver example.

Williams

The critical target of 'Moral Luck' is the conception of agency according to which the moral appraisal of our actions ought to depend only on what we put in, not on what comes out.⁷ Bernard Williams and Thomas Nagel, in separate works, attributed that conception to Kant but I will not be concerned with the accuracy of that attribution,⁸ as their chief point does not turn on it. That point, as I understand it, is that

⁷ I take this formulation from Gardner (2004) at 61.

⁸ For the record, I am sympathetic to Gardner's claim that Kant, the ostensible target of Williams's argument, generally allows luck to affect the moral assessment of a person, the only exception being the assessment of the virtue of a person of good will, *ibid.* at 62–63.

while the idea that people should be subject to moral appraisal only for things they can control (the ‘control principle’) holds much appeal, a considerable amount of our practical moral intuitions are inconsistent with it, so we have to revise one or the other.

Williams uses several, and now famous, examples to illustrate the significance of moral luck, but I want to focus on the one which has the nearest affinity to a problem of tort law: the example of a lorry driver who, through no fault of his, runs over a child. Williams begins by noting that the driver

will feel differently from any spectator, even a spectator next to him in the cab, except perhaps to the extent that the spectator takes on the thought that he himself might have prevented it, an agent’s thought. Doubtless, and rightly, people will try, in comforting him, to move the driver from this state of feeling, move him indeed from where he is to something more like the place of a spectator, but it is important that this is seen as something that should need to be done, and indeed some doubt would be felt about a driver who too blandly or readily moved to that position. We feel sorry for the driver, but that sentiment co-exists with, indeed presupposes, that there is something special about his relation to this happening, *something which cannot merely be eliminated by the consideration that it was not his fault*. It may be still more so in cases where agency is fuller than in such an accident, though still involuntary through ignorance.⁹

Here Williams argues that the driver’s agency remains related to the death of the child despite the fact that the driver was not at fault. However, he is not claiming that this entitles others to criticise the driver, or even that the driver ought to have any particular response to the accident. His claim, as I understand it, is twofold. The first part is diagnostic of the driver’s likely moral psychology after the accident. That diagnosis says that the driver will feel a special psychological connection to that accident, and a special sort of regret (‘agent-regret’) for having been involved as one of its causes. The second part is normative, and suggests not that agent-regret is due, but that the driver ought to not to rush the *transition* from their initial post-accident psychological state to the state of a mere spectator, who regards the accident as having no connection to their agency. That transition, Williams implies, has to be earned the hard way, and through the help of others.

Both claims strike me as eminently plausible, though the second could be straightforwardly accounted for in ways that do not require an appeal to moral luck. Williams is right to note that we would shudder if the driver, when asked for an explanation, said simply ‘hey, it wasn’t my fault, what’s for lunch?’ and went on as if the child’s death had absolutely nothing to do with him, morally speaking. But perhaps the ground of our intuition here is epistemic.¹⁰ When serious and life-changing harms are involved, we want to make sure that people have thought about the role in what happened without cutting corners, or putting themselves in the clear too easily.

⁹ Williams (1981) at 28 (my emphasis).

¹⁰ See also Enoch and Marmor (2007) at 415–418.

In certain situations, a failure to get more pairs of eyes on the case, or to elicit more perspectives on what happened and what could have been done, is a moral failure in itself, even if one turns out to have had adequate ground to consider oneself blameless. On that reading, the lorry driver's example becomes a lesson not in the sensitivity of our ideas about agency to matters of luck, but in the importance of epistemic circumspection.

Williams's discussion is even more guarded when it comes to a question closer to the tort lawyer's interests, that of institutional arrangements of compensation. He describes the various possibilities as follows:

The lorry-driver may act in some way which he hopes will constitute or at least symbolise some kind of recompense or restitution, and this will be an expression of his agent-regret. But the willingness to give compensation, even the recognition that one should give it, does not always express agent-regret, and the preparedness to compensate can present itself at very different levels of significance in these connexions. We may recognise the need to pay compensation for damage we involuntarily cause, and yet this recognition be of an external kind, accompanied only by a regret of a general kind, or by no regret at all. It may merely be that it would be unfair for the sufferer to bear the cost if there is an alternative, and there is an alternative to be found in the agent whose intentional activities produced the damage as a side-effect... A test of whether that is an agent's state of mind... is offered by the question whether from this point of view insurance cover would do at least as well. Imagine the premiums already paid (by someone else, we might add, if that helps to clarify the case): then if knowledge that the victim received insurance payments would settle any unease the agent feels, then it is for him an external case.¹¹

Note, again, that Williams does not claim that the driver *ought to* feel the famous feeling of 'agent-regret', or that he ought to experience the recognition of a need to compensate.¹² He certainly does not claim that the driver's recognition of that need ought to be of the 'internal' rather than the 'external' kind. The insurance hypothetical is not meant to identify which reaction is appropriate. Its sole purpose is to help distinguish between reactions on the basis of the driver's different possible psychological profiles. None of these claims is inconsistent with the control principle, because none of them involves a normative assessment.

The main point at which Williams signals that he wants to pick a fight over the best way to account for the lorry driver's agency appears a little further in the text:

What degree of such feeling is appropriate, and what attempts at reparative action or substitutes for it, are questions for particular cases, and that there is room in the area for irrational and self-punitive excess, no one is likely to deny. But equally it would be a kind of insanity never to experience sentiments

¹¹ Ibid. at 28–29.

¹² For the more ambitious view that treating ourselves as responsible for the outcomes of the exercise of our acting, rational agency is necessary for us to be able to understand ourselves as persons, rather than mere causes, see Honoré (1988), Perry (2001), and Gardner (2017) at 16–20.

of this kind towards anyone, and it would be an insane concept of rationality which insisted that a rational person never would.¹³

That claim rings true to me, and is sufficient to make Williams's point against the control principle: an account that took *any* experience of agent-regret to be an instance of irrationality would seem implausible for that reason. At the same time, this is not a claim about the attitude that an agent ought to have in relation to any particular accident or other harm-causing event. It is a claim about the agent's entire life. What Williams doubts is the sanity of the idea that a rational agent would *never* experience agent-regret. He does not suggest that such regret is always or even typically due, and makes no claim to the effect that the driver would display a moral virtue in offering compensation, or that others ought to encourage or make it easier for him to 'take responsibility' for the accident.

The explanation, I think, is that Williams is careful to distinguish the 'micro' question of what we are to make of a particular person in particular situations from the 'macro' question of what is the most plausible conception of our agency.¹⁴ It is at the macro level that Williams joins issue with the control principle. As he puts it in a Postscript to 'Moral Luck', he agrees that self-criticism is apt and due only when one could have avoided the bad outcome by choosing appropriately. His aim, rather, is to question

a presupposition... [that] can be put like this: as agents, we seek to be rational; to the extent that we are rational, we are concerned with our agency and its results to the extent that they can be shaped by our rational thought; to the extent that results of our agency could not be affected by greater rationality, we should regard them as like the results of someone else's agency or like a natural event. This idea seems to me very importantly wrong.¹⁵

This is an ethical claim, but it is of the macro type. It does not take a view about particular situations and involves no assessment of particular agents. Moreover, it allows that the connection between one's rational agency and its results may be strong enough for certain purposes, or along certain moral dimensions, but not for others. For example, the driver's causal involvement in the death may explain both why he might regard himself as having reason to *explain* himself to others, and why others may reasonably require him to supply such an explanation, e.g. to recount the situation so that others may form a view about whether or not he was at fault.¹⁶

¹³ Williams (1981) at 29.

¹⁴ I am using these unconventional terms to indicate that Williams's claim may be general, but it is still ethical rather than meta-ethical.

¹⁵ Williams (2010) at 245.

¹⁶ Cf. Schroeder (1997) at 357: 'Authors of actions have access to privileged information with respect to those actions that others lack. They can frequently reconstruct the practical reasoning that preceded the action, almost always better than external observers can. In so far as we must acknowledge a sense of authorship in order to sustain agency, the responsibility of discharging the burden of explanation that falls uniquely within an author's competence seems to suffice. Simply answering for one's actions by "owning up"—saying that is mine, and this is yours—permits each person to continue to construct and maintain an evolving personal history, which... is a prerequisite for personal identity and character'.

At the same time, the fact that causal involvement falls short as a justification for blaming the lorry driver or for holding him liable to pay compensation or to suffer criminal punishment suggests that in those contexts the threshold for finding an appropriate connection between the outcome and the driver's agency may be much higher. We should therefore expect accounts of the driver's agency to explain how the right connection between the driver and the accident may differ depending on the particular dimension of responsibility (self-narrative, blameworthiness, assignment of practical burdens) that we have in mind.¹⁷ Williams sees that his intuitions about moral luck may have traction only in some of those moral dimensions.¹⁸ His concern is simply that we do not take the fact that some of those dimensions require a close connection between a result and one's rational agency to justify the conclusion that such a connection is required in all of them.

Wolf and Enoch

Susan Wolf and David Enoch have taken Williams's thoughts about the role of luck in our understanding of agency as a basis for more specific ethical conclusions. In 'The Moral of Moral Luck', Wolf begins from Williams's intuition about the special connection that the lorry driver will feel towards the child's death, and imagines two drivers, who are both culpable for not having checked their brakes as thoroughly as they ought to have done before setting off for the day's work. The first driver runs over a child, the second does not, and the difference between their situations is wholly down to luck. The first driver experiences deep guilt and agent-regret about the accident, has nightmares over it, and feels the need to apologise and offer compensation to the family. The second driver, who reads about the first driver's accident in the paper, realises his good fortune, resolves to be checking the brakes thoroughly in the future, and otherwise goes about his life just as he did before. Wolf suggests, very plausibly, that 'if the first driver, who had actually killed the child, responded as I described the second driver doing, we would... be appalled and condemning'.¹⁹ She then describes the considerations that fuel this reaction in the following way:

The point is not that we expect the first driver to feel worse than we think he ought to feel, and that... third parties are required to bring his guilt feelings down to the appropriate level. Rather, it seems, we *want* the first driver to feel

¹⁷ Cf. Scanlon (2008). For an illuminating discussion of examples in which the agent, by dint of luck, has reason to endorse retrospectively choices that were clearly wrong and blameworthy at the time the agent made them (e.g. Williams's own Gauguin example), see Salow (2017).

¹⁸ I pass over the ways in which the choice of language can 'rig' an agent's self-narrative towards greater or lesser feelings of responsibility over an event. 'I killed the kid' and 'I was driving, a kid jumped in front of the car, and died' could figure in the driver's thoughts interchangeably, but each of them has different moral connotations. This suggests that the terms of the agent's self-narrative may not be a solid ground for conclusions about his responsibility, see also Wolf (2013) at 10.

¹⁹ Wolf (2013) at 9.

worse - he ought, at least initially, feel so bad that some soothing, some appeals to 'reason' are necessary to stop him from judging himself too harshly.²⁰

Wolf sees that, left at that, this claim would be liable to epistemic reduction.²¹ The reason why we want the first driver to feel worse at first and better only after we have thrashed out the details of the incident with him could be that this sequence is more likely to lead us to a proper examination of what happened. It stands to reason that a driver who brushed off the accident ('nothing to see here, folks') would be less likely to undertake a deep and critical examination of their role in it, or help others form an accurate view about the events. This also explains why, when the facts are finally in hand and the overwhelming influence of luck has been clearly established, it may be appropriate for the driver to feel a certain degree of relief. However, Wolf goes on to argue that such a reduction would fail to register a further feature of the situation. In experiencing the feelings in question, she argues, the driver would also be displaying a special moral virtue:

There is a virtue that I suspect we all dimly recognize and commend that may be expressed as the virtue of taking responsibility for one's actions and their consequences. It is, regrettably, a virtue with no name, and I am at a loss to suggest a name that would be helpful. It involves living with an expectation and a willingness to be held accountable for what one does, understanding the scope of 'what one does,' particularly when costs are involved, in an expansive rather than a narrow way. It is the virtue that... would lead one to apologize, rather than get defensive, if one unwittingly offended someone or hurt him. Perhaps this virtue is a piece or an aspect of a larger one which involves taking responsibility... for a larger range of circumstances that fall broadly within one's reach... It is not the case that the more responsibility one takes for the harms that lie at increasing distance from one's control, the better. Yet one ought to take responsibility for more than what, from a bystander's point of view, would be justly impersonally assigned.²²

As Wolf acknowledges, she is trying to thread a very fine needle. Saying that an agent 'took responsibility' or 'stepped up to the plate' does indeed signify that the agent did something good and worthy. However, the reasons underlying that favourable assessment may sometimes be of the garden-variety 'impersonal' kind. Saying sorry may be a good and worthy response not because it involves extending one's agency to take a responsibility that was not already one's own, but because it defuses a fraught situation and perhaps avoids a blame game. Sometimes an apology is virtuous because it serves as an apt prospective reminder, for oneself and for others, of the duties and responsibilities one already has, rather than because it expands one's agency to include outcomes that would otherwise fall outside its scope. Apologising can sometimes be good and worthy precisely because it preserves a degree of

²⁰ Ibid. at 10.

²¹ Ibid. at 7.

²² Wolf (2013) at 13.

ambiguity between being sorry that a state of affairs came about and being sorry for having brought it about. And we all know the type of person who thinks everything is their fault. Call them what you will, but they are not paragons of virtue.

At the same time, we can use cases where the assessment is based on impersonal reasons as a background against which to isolate those cases that instantiate the virtue that Wolf has in mind. Those seem to me to be cases in which the agent has causally contributed to a harmful outcome; that contribution does not make them responsible for that outcome on ‘impersonal’ grounds; the agent takes that contribution as giving them reason to undertake a further act or decision that forges a new connection between their agency and that outcome; this connection involves the agent treating themselves as responsible for that outcome; and the relevant act or decision counts as a display of virtue. The challenge Wolf sets herself is to explain when that last condition is met.

Her answer has two parts. The first involves an analogy from generosity. If it is virtuous to offer others more help or love or resources than is necessary or required by justice, then it is also virtuous to take on more burdens than one would be responsible for under the control principle.²³ This seems to me unsatisfactory, not least because it risks losing one of the very ideas that motivates Wolf’s argument, namely that the agent who takes responsibility does so in response to their *causal contribution* to the outcome. While both the generous person and the person who takes responsibility over a state of affairs think ‘I have to do something about this’, only the latter thinks ‘because I brought that state of affairs about’, and it is the force of this further thought that Wolf is trying to explain.²⁴ In fact, if the generous agent did have that further thought, we might not regard their action as generous at all. It follows that the appeal to generosity (or beneficence, charity, compassion, and similar virtues) cannot explain what, if anything, might be virtuous about taking responsibility.

The weight of Wolf’s case falls on the second argument, which she terms one of ‘psychic health’:

We expect the driver to offer to pay, then, not only because we want him to be generous, but because we expect him to accept contingency in the determination and assessment of who he is... The reason for objecting quite generally to an attitude of greater detachment and for commending an embrace of at least some of what lies beyond the sphere of one’s will has less to do with a benevolent concern for others than with a view about what, for lack of a better word, might be called *psychic health*. The desirability of this trait comes partly from its expressing our recognition that we are beings who are thoroughly in-the-world, in interaction with others whose movements and thoughts we cannot fully control, and whom we affect and are affected by accidentally as well as

²³ Wolf (2013) at 14.

²⁴ To put it in the famous terms of von Wright (1963), Chapter 3, the person who takes responsibility for an outcome regards that outcome as a ‘result’ of the exercise of their agency, whereas the generous person regards that outcome as a ‘consequence’ which they have reason to help improve or repair.

intentionally, involuntarily, unwittingly, inescapably, as well as voluntarily and deliberately.²⁵

David Enoch reaches a similar conclusion through a different path. Whereas Wolf sees taking responsibility as involving the display of moral virtue, he sees it as involving the exercise of a moral (or normative) power.²⁶ The idea that agents have the power to bring about changes to their normative profile is familiar mostly in the context of interpersonal relationships, but Enoch believes that it can be extended to the relationship between an agent and the consequences of their actions. And like Wolf, he believes that such an extension can be sufficiently grounded on the fact that we ‘take interest’ in the world and in the reason this gives us to incorporate at least some of the ‘unintended or unforeseen consequences’ of our actions into our agency.²⁷

Whether we approach taking responsibility as a moral virtue or a moral power (or both), accepting this line of argument may have implications for institutional action. For example, it is arguable that our institutions ought to encourage virtuous conduct, or at least ought to not make it difficult for agents to lead a life of minimal moral virtue.²⁸ Depending on the stringency of that duty, perhaps tort law ought to provide sufficient incentives to, or eliminate certain counter-incentives for, agents to ‘take responsibility’ over accidents that lie outside the narrow field of their agency. Alternatively, if we see taking responsibility as the exercise of a moral power then, as Enoch says, ‘there is some reason to plan legal arrangements that will encourage the exercise of this power’.²⁹

Both Wolf and Enoch note that their respective discussions of moral luck and taking responsibility are exploratory, and that the assessments they put forward have lots of moving parts. Their core claim is that we should make room in our moral and legal imagination for the idea that taking responsibility is an important and distinctive part of our ethical life, without committing themselves to specific claims about the strength of the reasons that this idea gives rise to, or its relation to other applicable ideas. In what follows, I will assume that Wolf’s and Enoch’s core claim is correct, and raise three doubts about their respective explanations of its truth. First, whether the decision to take responsibility counts as a display of virtue or as a justified use of the agent’s moral powers seems independent of the relation between that decision and the agent’s psychic health or their interest in engaging with the world. Second, it is not clear that giving institutional form to the virtue or power to take

²⁵ Wolf (2013) at 14 (my emphasis).

²⁶ Enoch (2014) at 266: ‘The power to take responsibility for things that lie outside the scope of one’s core agency is, it seems to me, a constitutive part of valuable relationships and ways of thinking of ourselves. As Gary Watson—following Joseph Raz—puts the point: normative powers “can only be justified if the creation of such special relationships between people is held to be valuable.”’, citing Watson (2009) at 162.

²⁷ *Ibid.* at 267–268.

²⁸ See e.g. Shiffrin (2007) at 717: ‘when a legal practice is pervasive and involves simultaneous participation in a moral relationship or practice, the content and normative justification for the legal practice must be acceptable to a reasonable moral agent with a coherent, stable, and unified personality’.

²⁹ Enoch (2014) at 268.

responsibility will have the moral benefits that Enoch envisages. Third, Wolf and Enoch may be missing that the relationship between our ideas about agency and our ideas about institutional design goes both ways: a well-designed system of liability may sometimes help us reconsider who we are and what we are responsible for.

The case for Wolf's and Enoch's main claim turns on the idea that hypothesising a virtue or power of 'taking' responsibility helps promote our interest in our psychic health and our engagement with the world. One problem with that claim is that whether that virtue or power is actually in display will turn on considerations that have little to do with either interest. In fact, improvement in psychic health or one's feeling of engagement with the world would seem to be at best epiphenomenal and at worst irrelevant to that question. The agent whose decision to take responsibility causes them to feel psychologically destroyed is no less virtuous than the agent who feels psychologically healthier for having taken that decision. Similarly, the conditions for the justified exercise of the moral power Enoch describes do not depend on whether the agent feels more or less engaged with the world as a result. If those conditions depend on the experiences, emotions, and other reactions of the agent who takes responsibility, they depend on the experiences, emotions, or reactions that this agent *has reason* to have. Wolf and Enoch might, of course, claim that acting on the relevant reasons will tend to promote one's psychic health, or the agent's engagement with the world, even if it does not have that effect on every occasion.³⁰ The problem is that this claim can be made reasonably easily for any situation in which the agent responds well to the applicable reasons, whatever those may be. Imagine an agent who has reflected deeply about their role in some serious accident and has concluded, correctly, that they are not responsible for it. This agent's psychic health is in no worse shape, and their engagement with the world is no weaker than that of the agent who decides to take responsibility for the accident. This suggests that perhaps what promotes the relevant interests is not the decision to take responsibility, but appropriate reflection on *whether* to take responsibility.

This seems to me to expose a further tension in Wolf's version of the argument. To defend the distinctive character of the virtue she describes, she claims that it does not reflect the agent's benevolent concern for others but the agent's deep interest in recognising that they are in-the-world, and their actions have 'effects' on others. However, the decision to take responsibility *itself* has effects on others, and so does the decision not to take responsibility. Establishing whether the agent who takes responsibility is doing something virtuous on the basis of its effects on others therefore risks collapsing to the question of what the agent has reason to do all things considered.³¹ Suppose that the lorry driver takes responsibility over the child's death

³⁰ David Enoch has suggested (in correspondence) that the story about engagement with the world speaks to why we have the relevant moral power, not to the conditions for its justified exercise (just like certain stories about the value of promising, e.g. that it promotes valuable relationships, need not be stories about the conditions for the validity of a promise). I agree that this two-track explanation is available to him in principle, but for the reason given in the text, I am not sure that it identifies a distinctive value in an agent's decision to take responsibility.

³¹ I hasten to add that even if grounding the virtue in question on the effects of the agent's conduct on others collapses into a proposition about what the agent has reason to do, not all propositions about

because he finds that essential to maintaining his psychic health, but that his gesture makes it less likely that anyone will establish what really happened and find out where the responsibility ought to lie (e.g. because all interested parties are happy to treat the case as closed). Given that assigning responsibility to the right person is something we have good reason to care about, the lorry driver's action would, amongst other things, obstruct the course of justice. We might, of course, treat this as a situation of conflict and say that the driver is doing something virtuous but that the requirements of justice present us with countervailing reasons to seek instead the person with whom responsibility ought to lie. My worry is that this view of the matter lets the driver off the moral hook, insofar as it excludes from the discussion of the driver's virtue the reasons he has to be mindful of how their decision to take responsibility will affect our investigation. The alternative would be to say that when such reasons are present and accessible to the driver, the driver's act may not be virtuous at all.³² But either way, by asking us to focus our attention on the agent's psychic health and their need to feel engaged with the world, Wolf may be constraining unduly our view of the virtue she is pointing to.

A second worry is that the benefits of casting the taking of responsibility in institutional form are not as clear as Enoch suggests. Say that a legal system does not make allowance for agents to take responsibilities which the law has assigned elsewhere. It would not follow that there is anything wrong with that system, morally speaking, as the agent may have sufficient opportunity to practise the relevant virtue or power outside the law and its institutions. After all, one may take responsibility for an outcome in many ways that do not involve or presuppose any institutional involvement, e.g. by offering a personal apology for having caused it, or by taking it upon oneself to look out for the victim in the future. Enoch's point is that the law can add value to the agent's exercise of the moral power to take responsibility, because it can assure the agent, the victim, and others that, once the agent has decided to take responsibility, they will not be able to backtrack at will. This can work well if the agent chooses some conventional form of responsibility-taking, such as paying the victim compensation or providing a certain service to them without payment. The law may then play its part by holding that the victim-recipient of that compensation or service has not been unjustly enriched, or that the agreement between the agent and the victim is exempt from contract rules about consideration and so on.³³ Enoch believes that something similar goes for apologising. In discussing the New Zealand system, which replaces individual tort liability for personal injury with a public insurance scheme, he argues that such a system could be rendered consistent with

Footnote 31 (continued)

virtue collapse in this way. For a recent defence against this Humean challenge to virtue ethics, see Korsgaard (2019) at 101–103.

³² Wolf (2013) at 15 hints at that possibility.

³³ The common law of unjust enrichment is already sensitive to Enoch's proposal. The agent who wants to claim back compensation they paid as a gesture of taking responsibility will be unable to show that the victim's enrichment was due to an 'unjust factor'. See Mitchell (2013) at 1034–1036.

the exercise of the moral power of taking responsibility by providing agents with an institutional route to apology.³⁴

Enoch may still be underestimating the risk that such institutional measures may end up overshooting their own target. The problem is not that having the state supervise apologies is wrong, but that it is hard to tailor the form of that supervision to the case for offering apologies *as a way of taking responsibility*. In line with Williams's view, Enoch recognises that taking responsibility is something that one will do occasionally or exceptionally, not as a matter of course. The sphere of one's responsibilities will be mostly populated by responsibilities one has, rather than by responsibilities one takes. Accordingly, their claim is not that one's psychic health or one's power to engage with the world improves the more responsibility one takes over the consequences of one's actions, but that something would be wrong with an agent who *never* took responsibility for anything. However, to fashion an agenda of institutional reform from this more measured 'macro' claim, one would have to explain how a system that facilitates the merely occasional taking of responsibility would ensure that agents are not tempted or pressured into overusing that facility, i.e. how it would ensure that the occasional will stay occasional, rather than become the new normal. The obvious risk is that casting the taking of responsibility in institutional form can contribute to the creation of dangerous heuristics and social practices: the next person, who is not feeling up to taking responsibility in the same way, may end up being judged under a false high standard. It is therefore important to design the relevant rules in ways that prevent agents' exposure to such pressure. The problem is that it is hard to suggest a legal design that could deliver on that score. The law is excellent in giving people directives and powers. To facilitate the occasional taking of responsibility, it would have to do another job too, and one which it seems awkwardly placed to undertake, namely to give people a power while ensuring that they are not exercising it *too often*. In the absence of concrete suggestions about how the law could manage this, it is arguable that systems that are not very prescriptive about when and how agents may take responsibility are already getting it fairly right.

A third and related worry is that perhaps taking responsibility over an event or outcome will be a display of virtue or a justified exercise of a normative power on the condition that there is no institutional alternative that deals adequately with that event or outcome. In particular, the fact that a society provides general insurance coverage against certain harms may change our assessment of an agent who attempts to take responsibility for having caused a harm of the relevant type. Here is an example. Safety features in modern cars are now better than ever, and have been quite successful in reducing accidents and injuries. Some of those features lower the stakes of poor driving choices, e.g. airbags reduce the risk of serious injury on impact, sensors warn drivers about the proximity of nearby cars when parking, and so on. Other features block or correct poor driving choices, e.g. cars do not start until the driver has fastened their seat belt, anti-collision systems kick in when the driver has failed to apply the brakes, and so on. All those systems save lives. They also require less of people by way of driving skill and, when fully self-driving cars

³⁴ Enoch (2014) at 270–271.

become a reality, they will not require people to have any driving skill at all. Overall, the average driver of the future will have less driving skill than the average driver of today, and much less skill than the driver of 20 years ago. Romantics will find something to mourn in this development, but few of them would claim that exposing people's lives to greater risk than necessary is a fair price for having more people exercise a higher level of driving skill. While driving skills remain valuable, and a good driver is still worthy of admiration, once technology allows people to move around safely without having to exercise a certain level of driving skill, we have reason to move away from forms of transportation that require people to do so. We think well of an agent who has those skills, and may even pay good money to watch the agent compete with others for sporting or entertainment purposes, but we do not think worse of an agent who lacks or does not even care to acquire them. Driving mastery becomes another laudable but optional pursuit, like being able to mend one's socks, or to light a fire by rubbing sticks together.

Perhaps something similar applies in respect of an agent's conception of themselves and their responsibilities. Wolf and Enoch are not alone in taking seriously our natural impulse to think of ourselves not simply as causes of consequences in the world, but also as the authors of the outcomes that are causally connected to the exercise of our rational agency.³⁵ When those outcomes are harmful, we think about what we could have done differently, we reconsider decisions and acts that looked perfectly benign at the time, and so on. Sometimes we also feel that it falls on us to pick up the pieces, even when morality or the law do not assign that responsibility to us. At the same time, we should avoid taking the impulse to regard ourselves as responsible in those circumstances as a ground to reverse-engineer a virtue out of the absence of institutional alternatives for dealing with those outcomes. Certain advantages of putting in place such alternatives, e.g. tax-funded comprehensive personal injury insurance coverage, are obvious. Victims of harm receive compensation regardless of the depth of the injurer's pockets, and injurers avoid the prospect of financial ruin for what is often momentary carelessness. But perhaps systems that socialise the cost of accidents have the further advantage of relieving the pressure people might feel to take individual responsibility for accidents, insofar as it ensures that the cost of those accidents does not fall on persons who might be poorly placed to bear it. When such a system is in place, taking a responsibility that is not already one's own may be no different than acquiring a wholly optional skill, and should not be prized any higher.

This view need not involve a withdrawal from the world, or a failure to recognise that 'we are beings who are thoroughly in-the-world, in interaction with others whose movements and thoughts we cannot fully control'. Taking on responsibilities that are not already ours is one way of being in the world, but it is not the only one, and it may

³⁵ See n.11 above. Raz (2010) at 17 draws a more limited conclusion from the idea that responsibility 'is crucial to our sense of being in the world', namely that we are responsible for the successful or failed exercises of our powers of rational agency, but the difference is not germane to my present purposes.

not even be the best.³⁶ We can be in the world precisely *by* taking a more modest view of the reach of our agency, and using our moral energy and imagination to devise better ways to keep ourselves and others safe and improve the world we share. This is the way we think about safety standards. If we can make an activity safer, use technology to reduce the cost of its going wrong, or spread that cost amongst the community of people who engage in or benefit from that activity, we have reason to do just that. The older safety standards may have required people to exercise more skill and judgement, but that is not sufficient reason for us to yearn for a return to those standards, and it is not often that such yearning can withstand close scrutiny. The challenge for Wolf and Enoch is to explain what might be wrong with thinking about the relationship between public insurance schemes and individual agency in the same way.

Conclusion

Williams drew our attention to what might be wrong with denying the role of luck in our understanding of agency and responsibility. Wolf and Enoch have asked us to focus instead on what might be virtuous and valuable in embracing that role, and on how our institutions might assist us in that regard. They have argued that the agent who ‘takes’ a responsibility that the law or morality does not already assign to them may be displaying a special moral virtue or exercising a distinctive moral power, and they have accounted for that virtue or power in terms of the way taking responsibility enhances the agent’s sense of being-in-the-world. While the argument seems attractive to me, I have raised certain doubts about the case for it, and Enoch’s particular attempt to extract institutional implications from it. The deeper of those doubts is that what conduct counts as virtuous or as a justified exercise of a normative power will sometimes depend on the design of the relevant institutional structures. Sometimes our institutions do not help us simply to ‘act out’ our ideas about what actions or events we are responsible for, but also to reconsider those ideas by giving us good practical alternatives for dealing with the outcomes of those actions or events. Just like a good safety measure removes the need for the exercise of individual skill in avoiding accidents, a good system of liability may remove the need for people to take responsibility for such accidents when they do happen. If such safety measures do not jeopardise our sense of being-in-the-world, neither do such systems of liability. This seems to me something that Williams missed in his original suggestion about the effect of public liability insurance on the lorry driver’s attitude towards the child’s death. Williams asked us to assume that the premiums have been paid, and claimed that if the lorry driver does not experience agent-regret he would be treating the accident as an event ‘external’ to his agency. I hope to have shown that there is a more complex possibility. The driver who does not experience agent-regret may still feel that the accident is internal to a certain aspect of his

³⁶ I am not claiming that institutional design has metaphysical priority over our ideas about agency, only that it may affect the application of those ideas to individual action. I am particularly grateful to David Enoch for discussion on that point.

agency: the aspect constituted by his membership in a community that gives injurers and victims a measure of protection against the vagaries of bad luck.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Enoch, D. 2012. Being Responsible, Taking Responsibility, and Penumbral Agency. In *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams*, ed. U. Heuer and G. Lang, 95. Oxford: Oxford University Press.
- Enoch, D. 2014. Tort Liability and Taking Responsibility. In *Philosophical Foundations of the Law of Torts*, ed. J. Oberdiek, 250. Oxford: Oxford University Press.
- Enoch, D., and A. Marmor. 2007. The Case against Moral Luck. *Law and Philosophy* 26: 405.
- Gardner, J. 2004. The Wrongdoing that Gets Results. *Philosophical Perspectives* 18: 53.
- Gardner, J. 2017. The Negligence Standard: Political not Metaphysical. *Modern Law Review* 80: 1.
- Goldberg, J., and B. Zipursky. 2007. Tort Law and Moral Luck. *Cornell Law Review* 92: 1123.
- Honoré, T. 1988. Responsibility and Luck: The Moral Basis of Strict Liability. *Law Quarterly Review* 104: 530.
- Jackson, C. 2012. Tort Law, Moral Luck, and Blame. *Cleveland State Law Review* 60: 57.
- Keating, G. 2006. Strict Liability and the Mitigation of Moral Luck. *Journal of Ethics & Social Philosophy* 2: 1.
- Korsgaard, C. 2019. Constitutivism and the Virtues. *Philosophical Explorations* 22: 98.
- Mitchell, C. 2013. Unjust Enrichment. In *English Private Law*, 3rd edn, ed. A. Burrows, 1031. Oxford: Oxford University Press.
- Perry, S. 2001. Honore on Responsibility for Outcomes. In *Relating to Responsibility: Essays for Tony Honore*, ed. P. Cane and J. Gardner, 71. Oxford: Hart Publishing.
- Raz, J. 2010. Responsibility and the Negligence Standard. *Oxford Journal of Legal Studies* 30: 1.
- Salow, B. 2017. Partiality and Retrospective Justification. *Philosophy & Public Affairs* 45: 8.
- Scanlon, T. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Harvard University Press.
- Schroeder, C. 1997. Causation, Compensation, and Moral Responsibility. In *Philosophical Foundations of Tort Law*, ed. D. Owen, 347. Oxford: Oxford University Press.
- Shiffrin, S. 2007. The Divergence of Contract and Promise. *Harvard Law Review* 120: 708.
- von Wright, G. H. 1963. *Norm and Action*. Abingdon: Routledge & Paul.
- Waldron, J. 1995. Moments of Carelessness and Massive Loss. In *Philosophical Foundations of Tort Law*, ed. D. Owen, 387. Oxford: Oxford University Press.
- Watson, G. 2009. Promises, Reasons, and Normative Powers. In *Reasons for Action*, ed. D. Sobel and S. Wall, 155. Cambridge: Cambridge University Press.
- Williams, B. 1981. Moral Luck. *Moral Luck: Philosophical Papers 1973–1980*, 20. Cambridge: Cambridge University Press.
- Williams, B. 2010. Moral Luck: A Postscript. *Making Sense of Humanity, and Other Philosophical Papers 1982–1993*, 245. Cambridge: Cambridge University Press.
- Wolf, S. 2013. The Moral of Moral Luck. *Philosophic Exchange* 31: 5. Reprinted in Calhoun, C. 2014. *Setting the Moral Compass: Essays by Women Philosophers*, 113. Oxford: Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.