

# Accountability and Intervening Agency: An Asymmetry between Upstream and Downstream Actors

Saba Bazargan-Forward

UC San Diego

## Abstract

Suppose someone (P1) does something that is wrongful only in virtue of the risk that it will enable another person (P2) to commit a wrongdoing. Suppose further that P1's conduct does indeed turn out to enable P2's wrongdoing. The resulting wrong is agentially mediated: P1 is an enabling agent and P2 is an intervening agent. Whereas the literature on intervening agency focuses on whether P2's status as an intervening agent makes P1's conduct less bad, I turn this issue on its head by investigating whether P1's status as an enabling agent makes P2's conduct more bad. I argue that it does: P2 wrongs not just the victims of  $\phi$  but P1 as well, by acting in a way that wrongfully makes P1 accountable for  $\phi$ . This has serious implications for compensatory and defensive liability in cases of agentially mediated wrongs.

## 1

When one person undertakes conduct that is wrong only in virtue of the risk that it will enable some other person's wrongdoing, and that conduct does indeed turn out to enable that wrongdoing, then the resulting wrong is agentially mediated. Here are two examples.

### 'Race'

P2 would like to participate in a reckless and illegal drag-race, but her car is in the shop. P1 agrees to allow P2 to borrow P1's car. P1 is aware that P2 plans on using it to race. During the race, P2 strikes an uninvolved pedestrian, causing serious injuries to her. If it were not for P1's assistance, P2 would not have been able to

participate, and consequently the pedestrian would not have been injured.

In this example P1 plays an indispensable role in causing the pedestrian's injuries. But the injuries were committed by P2 – not by P1. We can say that the harm which P1 brought about was 'agentially mediated' by P2, and that P2 was an 'intervening agent'.

In Race, P1 knows that agreeing to P2's wishes facilitates P2's intended wrongful actions. But agential mediation includes cases in which P1 does not foresee what P2 will do, though P1 is in a position to foresee it:

'Bulldozer 1'

The employee of a construction company (P1) negligently leaves keys in a bulldozer overnight. The risk that makes this negligent manifests; a local vandal (P2) gains unauthorized access to the bulldozer, taking it for a joy-ride recklessly causing property damage.

In this case, P1 is in a position to know that what he was doing risked unauthorized access to the bulldozer. This too is an example of agential mediation, which can be described more generally as follows. A harm  $\phi$  is agentially mediated by P2 with respect to P1, act A1, act A2, for times t1 and t2, if:

- (i) P1 commits A1 at t1, which significantly increases the probability that P2 will voluntarily commit A2 at t2, which causes a harm  $\phi$ .
- (ii) P2 would not have committed A2, and  $\phi$  would not have occurred, if P1 had not committed A1.
- (iii) P1 knows (or is in a position to know) both of the above.

On counterfactual and probability-raising theories of causation, what P1 does would probably count as causally contributing to P2's wrongful act. But on causal process theories of causation, such as mark-transfer theories and conserved-quantity-transfer theories, what P1 does might not count as causally contributing to P2's wrongful act. Still, even on these theories, the fact that P1's act is a *sine qua non* of  $\phi$  provides a prima facie moral reason for P1 to curb her conduct accordingly. And it provides a prima facie basis for regarding P1 as at least somewhat morally responsible for what P2 does

(without thereby diminishing P2's responsibility for her own conduct). To remain neutral between competing theories of causation, I will not say that P1's act *causally* contributes to  $\phi$ . Instead, I will stipulate that P1's act contributes to  $\phi$  in so far as that act is a *sine qua non* of  $\phi$ . (This will, of course, leave unanalysed cases of pre-emption and over-determination, but here I am concerned only with the simplest and most straightforward case of intervening agency, which, as we shall see, is challenging enough.)

Compatibilists and incompatibilists of nearly all stripes who countenance attributive responsibility can accept the foregoing claims. This might not seem true of libertarianism, but nearly all libertarians believe that P1 can causally influence the factors that P2 takes into consideration when deciding what to do, thereby indirectly affecting (though not determining) P2's decision to commit A2. This leaves P2 room to exercise sovereignty over the decision to commit A2, while causally implicating P1 in that decision. On radical forms of libertarianism, however, nothing external to the freely willed decision to act can causally influence that decision (not even one's own character). But not many hold this view.<sup>1</sup>

To avoid confounding variables, it is important to keep morally relevant features constant between P1 and P2. I will assume that though P1 intends A1 and P2 intends A2, *neither intends*  $\phi$ . I will also assume that P1 is in a position to recognize that committing A1 imposes a risk of  $\phi$ 's occurrence via P2's agency; likewise, P2 is in a position to recognize that committing A2 imposes a risk of  $\phi$ 's occurrence.

A final preliminary point: there are at least two kinds of intervening agency. P1 *enables* P2 by providing P2 with otherwise absent means of committing A2, which causes  $\phi$ . But  $\phi$  also counts as agentially mediated if P1 *motivates* P2 by drawing her attention to motivating reasons to commit A2. I suspect that whether P2 is motivated or enabled is morally relevant to the evaluation of P1's and P2's conduct. But here I am concerned with agential mediation *simpliciter* – I do not distinguish between enabling and motivating, which I leave to future investigations.

---

<sup>1</sup> This view was held by Duns Scotus (J. F. Ross and T. Bates, 'Duns Scotus on Natural Theology', *The Cambridge Companion to Duns Scotus*, ed. T. Williams (New York, 2002), pp. 193–237, at 222) and by Jean-Paul Sartre (*Being and Nothingness* (Washington, 1993), pp. 567–8). They would deny that P1 contributes to  $\phi$ .

Given this account of intervening agency, is agential mediation relevant to how we should assess P1's and P2's conduct? If so, in what way? The existing literature on the moral relevance of intervening agency is dominated by a narrower version of that question: does P2's status as an intervening agent diminish P1's responsibility for  $\phi$ ?<sup>2</sup> A prominent pre-theoretic view is that when the upshot of one's action 'passes through' the agency of another, one's responsibility for the upshot is diminished or even eliminated – as if the other person were like a resistor in an electrical circuit. As Sanford Kadish put it, we have 'an ingrained view of the world' according to which 'the causal route through which an upshot occurs is a central feature in assessing blame'. Consequently, 'recklessness with respect to a natural happening is not seen to be commensurable with recklessness with respect to another person acting in a certain way'. He confesses, though, that '[w]here this view comes from is a bit mysterious'.<sup>3</sup> To borrow a phrase from Samuel Scheffler, the notion that intervening agency is morally relevant in this way 'remains an intuition in search of a foundation'.<sup>4</sup>

The 'Voluntary Intervention Principle' (VIP), versions of which are accepted in Anglo-American criminal and tort law, reflect that intuition in so far as the VIP states, in the words of Hart and Honoré, that: 'The free, deliberate and informed intervention of a second person, not acting in concert with the first, and intending to bring about the harm which in fact occurs or recklessly courting it, is normally held to relieve the first actor of criminal responsibility.'<sup>5</sup> Hart and Honoré attempt to defend this view by appealing to an agent-causal view of human action. I will not comment on the

---

<sup>2</sup> See, for example H. L. A. Hart and Tony Honore', *Causation in the Law*, 2nd edn. (Oxford, 1958); Sanford Kadish, 'Complicity, Cause and Blame', *California Law Review* 73 (1985), pp. 323–410; Michael Zimmerman, 'Intervening Agents and Moral Responsibility', *The Philosophical Quarterly* 35 (1985), pp. 347–58; Heidi Hurd, 'Is it Wrong to Do Right When Others Do Wrong?', *Legal Theory* 7 (2001), pp. 307–40; and Michael S. Moore (M. S. Moore, *Causation and Responsibility* (New York 2009).

<sup>3</sup> Sanford H. Kadish, 'Reckless Complicity', *The Journal of Criminal Law and Criminology* (1997), p. 393.

<sup>4</sup> Samuel Scheffler, *The Rejection of Consequentialism* (Oxford, 1982), p. 112

<sup>5</sup> Hart and Honoré, *Causation in the Law*, pp. 42-3. Kadish similarly argues that in criminal law 'voluntary actions cannot be said to be caused' (Kadish, 'Complicity, Cause and Blame', p. 371).

(im)plausibility of this defence, as it has been sufficiently criticized by others.<sup>6</sup>

The focus on whether intervening agency should affect our assessment of P1's conduct has obscured a distinct and equally important issue which has gone unnoticed: whether intervening agency should affect our assessment of P2's conduct. This is the question I will be focusing on here. That is, I will not be investigating whether P2's status as an intervening agent *makes P1's conduct less bad*. Instead, I will turn this issue on its head by investigating whether P2's status as an intervening agent *makes P2's conduct worse*. And I will argue that it does: holding all else fixed, P2 is responsible for *two* wrongs whereas P1 is responsible for one. Each of P1 and P2 bears some responsibility for  $\phi$ . But in addition, P2 wrongs P1 by having wrongfully made it so that  $\phi$  is a consequence of what P1 does. I will argue that this has important implications for the comparative liability of P1 and P2 for  $\phi$ .

An advantage of my argument is that it grounds the claim that P2 acts more wrongfully than P1 without having to advert to a pre-theoretic, 'ingrained view of the world' according to which there is an intrinsic morally relevant difference between *causing* a harm and *enabling* it. There are, instead, principled reasons for thinking that agential mediation is morally relevant to our comparative moral assessments of P1 and P2.

But before presenting my account, I must briefly explore the relevance of probability to assessing the comparative responsibility of P1 and P2 for  $\phi$ . One might immediately point out that there is already on offer a rather mundane reason for thinking that P2's conduct will be worse than P1's. All things being equal, someone who imposes a greater risk of wrongful harm has acted in a more objectionable way than someone who imposes a lesser risk of wrongful harm. So if we can show that P2 imposes a greater risk of wrongful harm than P1, then we have thereby shown that P2 acts in a more objectionable way than P1, *ceteris paribus*.

It might seem that we can indeed show that what P2 does is more risky than what P1 does, where P1 commits A1, thereby enabling P2 to commit A2,

---

<sup>6</sup> See especially Michael S. Moore, 'The Metaphysics of Causal Intervention', *California Law Review* 88 (2000), pp. 827–78

which in turn causes the harm  $\phi$ . When P1 commits A1, she is less likely to bring about the harm  $\phi$  than when P2 commits A2 since in the examples under consideration it is only via P2 that P1 brings about the harm. Committing A1 increases the risk that P2 will commit A2 by  $n1\%$  (where  $0 < n1 < 100$ ), which in turn increases the risk of  $\phi$ 's occurrence by  $n2\%$  (where  $0 < n2 < 100$ ). Hence P1 imposes a risk of  $n1\% \times n2\%$  that  $\phi$  will occur, which is less than the degree to which P2 raises the risk of  $\phi$ 's occurrence, since  $(n1 \times n2) < n2$ . This is just to note that the effect of any (token) cause is always less than fully certain; this uncertainty ramifies as we project forward to the effects of *those* effects. The result is that we are less certain that committing A1 will bring about  $\phi$  than we are that committing A2 will do so. Thus P1 acts in a less risky way when she commits A1 than P2 does when she commits A2. (Kadish seems to note similarly that intervening agency is relevant to probability when he writes that 'the intervention of a second actor, whose action is required for the harm to occur, reduces the probabilities that the harm would ultimately happen'.)<sup>7</sup>

This reasoning partly explains why P2's conduct in Race and Bulldozer 1 is more objectionable than P1's – *mutatis mutandis* for forthcoming cases. But it is not always the case that what P2 does is more risky than what P1 does. Consider this example.<sup>8</sup> Suppose that if P1 commits A1, there's a 95% chance that P3 will commit A2, and a 5% chance that P2 will commit A2. And if P3 commits A2, there's a 95% chance that the harm,  $\phi$ , will occur. But if P2 commits A2, then there is only a 5% chance that  $\phi$  will occur. Thus P1's act creates a 90.25% chance that  $\phi$  will occur through P3, and a 0.0025% chance that it will occur through P2. Hence P1's act is more risky than P2. In cases like these, advertent differences in risk-imposition cannot explain why we might think that what P2 does is worse.

The lesson here is that a probabilistic account of the moral difference between what P1 does and what P2 does tells only part of the story. Even when it is the case that what P2 does is more risky than what P1 does, there are additional, compounding reasons for thinking that what P2 does is worse than what P1 does. These reasons derive not from difference in risk-imposition but from the moral relevance of intervening agency *per se*.

---

<sup>7</sup> Kadish, 'Reckless Complicity', p. 381

<sup>8</sup> I owe this example to a referee at *Utilitas*.

## 2

I will argue that what P2 does is worse than what P1 does (for reasons independent of and in addition to their difference in risk-imposition). Specifically, I will argue that what P2 does is worse in that *P2 wrongs P1* but not vice versa. To show this, I will begin by trying to show that P1 has a claim against P2 that P2 refrain from committing A2, i.e. the act causally proximate to  $\phi$ .

By hypothesis, both P1 and P2 increase the probability of  $\phi$ 's occurrence. P1 does so by committing A1 at t1, and P2 does so by committing A2 at t2. But there is a further symmetry between the two actors. When each of P1 and P2 increases the probability of  $\phi$ 's occurrence, each thereby increases the probability that the other person will contribute to  $\phi$ . By committing A1 at t1, P1 increases the probability of  $\phi$ 's occurrence by increasing the probability that P2 will commit A2. Likewise, by committing A2 at t2, P2 increases the probability of  $\phi$ 's occurrence by increasing the probability that what P1 has done at t1 will contribute to  $\phi$ . To see this, consider again Bulldozer 1: A1 is the act of leaving the keys in the bulldozer, A2 is the act of stealing the bulldozer, and  $\phi$  is the harm of recklessly caused property damage. The employee, by leaving the keys in the bulldozer, increases the probability of unauthorized access to the vehicle, by providing an opportunity for the vandal to operate the bulldozer. And by taking advantage of the opportunity afforded by the employee's negligence, the vandal thereby ensures that what the employee has done will contribute to the bulldozer's theft.

But this symmetry between P1 and P2 conceals a fundamental asymmetry between them. At t2, P2 is in a position to *make it the case* that what P1 has done will contribute to A2. There is no point in time, however, at which P1 has this sort of influence over what P2 has done. In Bulldozer 1, at t2, the vandal can determine whether what the employee has done will contribute to the unauthorized access of the bulldozer. The vandal can do this simply by choosing to take advantage of the opportunity that the employee has unwittingly offered. The employee, however, is never in an analogous position. He can *enable* her, and might (if his mistake is one of recklessness rather than negligence) recognize the likelihood that the vandal

will consequently steal the bulldozer. But he is not in a position to *make* the vandal steal the bulldozer. So it is perfectly intelligible for the employee, in confronting the vandal after she is caught, to say ‘your actions made me a contributor to the theft of my employer’s bulldozer!’ But the vandal cannot felicitously say to the employee ‘your actions made me steal the bulldozer!’

The employee does not want the vandal to commit A2 – that is, the employee does not want the vandal to steal the bulldozer – not merely because the employee does not want the bulldozer stolen, *but because the employee does not want to have contributed to the bulldozer’s theft.* The vandal can, at t2, make it the case that what the employee has done will contribute to the bulldozer’s theft simply by stealing the bulldozer. By influencing the consequences of what someone else does, one influences what that person has done. Contrast this with the sort of influence that the employee has over the vandal. By leaving the keys behind, he can provide her only with the opportunity to steal the bulldozer – which is something that the vandal wants.

The crucial asymmetry between P1 and P2, then, can be summarized as follows: P2, in effect, co-opts P1’s act, by using it as a means to ends which P1 does not want. Imagine that the employee suddenly realizes that he left the keys in the bulldozer. He looks to a closed-circuit video feed, and sees on the screen the vandal approach the bulldozer. He watches in apprehension, knowing that what she does next will determine whether what *he* has negligently done will contribute to the unauthorized access of the bulldozer. At this moment, the employee is at the mercy of the vandal.

The suggestion here is that P2 is in a more privileged position than P1, since at t1, P1 can merely influence the probability that P2 will cause A2, whereas at t2, P2 is in a position to make it the case that P1 contributes to A2, by using the upshot of P1’s act as a means in furtherance of an end that P1 does not want.

At first glance, this might seem to get things backwards. After all, there is a sense in which P2 is standing on P1’s shoulders. Imagine that, at t1, the vandal from her hideout watches the employee close shop for the night. She watches in anticipation hoping that the employee leaves the keys behind. At that moment, the employee is in a position to determine whether the vandal can access the bulldozer – simply by refraining from leaving the keys behind.



In this sense, the vandal depends on the employee. Put more generally, were it not for P1, P2 would be unable to commit A2. But it is crucial to recognize that P1 is also standing on P2's shoulders – were it not for P2, A1 would not contribute to A2 (and therefore to  $\phi$ ).

The bare fact that P2 makes it the case that P1 contributes to  $\phi$  against P1's will does not diminish the badness of what P1 does. After all, adventitious events (i.e. events caused by happenstance) can similarly make it the case that P1 contributes to  $\phi$  against her will. Compare the following case with Bulldozer 1:

'Bulldozer 2'

The employee of a construction company negligently forgets to lock the parking brake of a bulldozer overnight when gale-force winds are expected. The risk that makes this negligent manifests; the wind pushes the bulldozer backward causing property damage.

In Bulldozer 1, P1 imposes an agentially mediated harm, whereas in Bulldozer 2 the harm he imposes is not agentially mediated. In both examples, an external event made it so that the employee indirectly imposed a harm that he would very much like to have avoided.

There is nonetheless a morally relevant difference between Bulldozer 1 and Bulldozer 2. In Bulldozer 1, unlike Bulldozer 2, the employee is *wronged*. When P2 makes it so that what P1 has done contributes to a wrongful end against P1's will, P1 thereby becomes accountable in so far as she becomes liable to proportionate compensatory or preventive harms.

The concept 'accountability' is a species of the polysemous concept 'responsibility': an individual is accountable for a harm just in case she is the proper subject of preventive or rectificatory liability. This is to say that we can impose deprivations upon her in order to prevent the harm or compensate its victims, without thereby wronging the accountable party. In the language of rights, the liable party has forfeited her right against being subjected to necessary preventive or compensatory deprivations. So as I use the concept, *accountability* implies *liability*. To say that an individual is morally liable *to* a harm is to say that she has forfeited her right not to be harmed in that way in furtherance of preventing or rectifying the harm *for which* she is liable. And the harm *for which* an individual is liable is the harm for which she is responsible.

So if for example the vandal escapes then the victims of the property damage that the vandal causes have a moral right to demand compensation from the negligent employee. Or if it turns out that the only way to prevent the damage is to harm the employee then she might be liable to be so harmed (depending on that harm's severity). Either way, P1, through her own negligence, has put herself in a position where she has a duty to compensate the victims of  $\phi$  or to prevent their victimization; *but she has this duty only if P2 commits A2*. The vandal, at t2, is in a position to determine whether what the employee did at t1 will serve as a basis for an enforceable duty against her.

Now, P1 has a legitimate interest in not being made potentially liable to preventive or compensatory harms, which is exactly what will happen if P2 commits A2. But the claim P1 has against P2 is not simply that P2 refrain from doing whatever would make P1 liable. It is unlikely that we have such a broad, free-standing claim against others. Rather, the sort of claim that P1 has against P2 is more complex: we have a claim against others that they refrain from using us by co-opting our actions in furtherance of wrongful ends which we repudiate, where pursuing or satisfying those ends would thereby make us liable. So P1's claim against P2 is not simply a claim that P2 refrain from acting in a way that will set back P1's interests, but that P2 refrain from *using* P1 in a way that will set back P1's interests.

But what if it turns out that P1 gets lucky in that there are no costs we can impose upon her to prevent or compensate the harm that she enabled P2 to commit? Consider the following case:

'Murder'

At t1, an assassin (P2) purchases a rifle from the only available seller (P1), who suspects that the assassin will use the firearm for illicit purposes. The seller is correct – the assassin intends to use the rifle to kill a politician she hates. She has only one chance to do so and it will require taking a difficult shot. In position with the rifle at t2, she takes the shot but misses, instead killing a bystander unintentionally. The bystander has no estate or friends morally entitled to compensation. And the bystander's murder could not have been prevented by imposing a harm on P1 (subsequent to having lent P2 the gun).

In this sort of case, P1 *would* be morally liable to compensatory harms if there were someone to compensate, and P1 *would* be morally liable to preventive harm if harming her were an effective means of preventing P2's act. But as it turns out, P1 is not actually liable in either way. It might seem, then, that P2 has not wronged P1 by acquiring the gun. This suggests that the principle of wrongfully imposed accountability is mistaken, or at least incomplete.

To address this challenge, it is helpful to say more about the nature of moral liability and the role it plays in our moral economy. Moral liability is necessarily instrumental in that a person is liable to a harm only if imposing that harm achieves a relevant good; this is because the purpose of liability is to prevent a rights-transgression or to restore a right.<sup>9</sup> When a person transgresses another's rights, the transgressor herself forfeits her right not be used as a means to preventing or rectifying the harms which she was transgressing. A rights-transgressor accordingly becomes liable to the (proportionate) means required to prevent or rectify that rights-transgression. So the fact that P1 is not liable in the case described above is wholly contingent – it just so happens that harming her does not have preventive or compensatory value.

It is useful, then, to distinguish between *actual* liability and *potential* liability.<sup>10</sup> An individual is merely potentially liable for a rights-transgression if that individual bears some responsibility for it, but imposing a proportionate harm on that individual would not prevent that rights-

---

<sup>9</sup> This account of liability is developed by Jeff McMahan. See especially: Jeff McMahan, 'The Basis of Moral Liability to Defensive Killing', *Philosophical Issues* 15 (2005), pp. 386–405; McMahan, *Killing in War* (New York, 2009); McMahan, 'The Conditions of Liability to Preventive Attack', *Gathering Threats: The Ethics of Preventive War*, ed. Deen K. Chatterjee (Cambridge, 2012), pp. 121–44. It is disputed, though, whether necessity is internal to liability, i.e. whether an individual can be liable to a harm even if there are less harmful means to preventing the harm for which the agent is liable. For discussion, see Jonathan Quong and Joanna Firth (J. Quong and J. Firth, 'Necessity, Moral Liability, and Defensive Harm', *Law and Philosophy* 31 (2012), pp. 673–701). Helen Frowe in particular argues that necessity is not internal to liability (Helen Frowe, *Defensive Killing* (Oxford, 2015), pp. 91–4).

<sup>10</sup> McMahan developed the notion of potential liability (which can be usefully distinguished from actual liability). See McMahan, *Killing in War*, pp. 21–2.

transgression or restore the transgressed right. A person is *actually* liable for a rights-transgression if she bears some responsibility for it *and* if imposing a proportionate harm on her will prevent the rights-transgression or restore the transgressed right.

It might seem that only actual liability is detrimental to us. That is why it might seem that P1 has not suffered a setback to her interests by P2 in Murder. After all, if P1 is merely potentially liable for a harm then no preventive or compensatory harms can be permissibly imposed on her; hence she has nothing to fear. But this view relies on a myopic understanding of our interests. There is a perfectly straightforward way in which P1 has been morally compromised by being made potentially liable. Knowing that others *would* be entitled to harm her if only doing so would benefit the victims of  $\phi$  makes her *usable* in a way antithetical to her interests – even if it turns out that she knows that she will never be so used. This is not to claim that a risk of a harm is itself a harm; rather, the claim is that P2 diminishes P1's moral status by making it permissible to harm her as a means if doing so will prevent or rectify the rights-transgression suffered by the victims of  $\phi$ . In this sense, being made *potentially* liable constitutes a setback to legitimate interests – namely, the interest in having our full moral status as inviolable beings intact.

Thomas Nagel makes this point in a discussion on the value of inviolability. He says: '[w]hat actually happens to us is not the only thing we care about: What *may* be done to us is also important, quite apart from whether or not it *is* done to us'.<sup>11</sup> Likewise, the 'status of heightened (logical) susceptibility to injury is, in a sense, a good thing to have, quite apart from any consequences or the likelihood of such injury'.<sup>12</sup> For this reason, being merely *potentially* liable is a setback to P1's interests – the bare fact that it makes her *usable* morally diminishes her.

There is an additional reason for thinking that being made potentially liable is a cost to P1. An individual who is potentially liable for a harm does not satisfy the circumstantial criteria for liability in that imposing an ostensibly preventive or compensatory deprivation upon her will not achieve any good.

---

<sup>11</sup> Thomas Nagel, 'The Value of Inviolability', *Morality and Self-Interest*, ed. Paul Bloomfield (Oxford, 2007), pp. 102–16, at 108.

<sup>12</sup> Nagel, 'The Value of Inviolability', p. 112.

However, someone who is potentially liable for a harm does indeed satisfy the agential criteria for liability – she bears sufficient responsibility for the harm to be the proper subject of proportionate preventive or compensatory deprivation if imposing such deprivations will be effective.<sup>13</sup> Now, satisfying the agential criteria for responsibility – especially when the individual does so by acting culpably – is enough to be the subject of reactive moral disapprobation from the community at large (or to at least be disposed to be subjected to such disapprobation). This is because it is only by dint of circumstantial moral luck that it so happens that the potentially liable individual is not actually liable. (Of course, if P2 refrains from committing A2, then P1 is the beneficiary of causal moral luck – but at least in such a case there is no harm for which to be the subject of moral disapprobation). So by committing A2, P2 imposes a cost on P1 in that P1 makes P2 responsible for a wrongful harm, which is something we all have in-principle reason to avoid, not only for the attendant negative social affects but also because being responsible for a wrongful harm is bad in itself. P1 does indeed, then, have a complaint against P2, since at t2 what P2 does *eo ipso* makes P1 responsible for  $\phi$ . P2 has no analogous complaint against P1 since P1 does not do anything at any point that *eo ipso* makes P2 responsible for  $\phi$ .

Against this, one might argue that P1 bears some responsibility for being wronged by P2. After all, P1 could have prevented this wrong simply by refraining from committing A1. But this does not diminish P2's wrong of making P1 impermissibly liable. As an analogy, suppose that J lives in a neighbourhood plagued with muggings. Tired of cowering in fear at home, J decides to go to an ice-cream parlour, despite knowing that there is a very high probability that she will be mugged. She does this as an act of defiance against the gangs plaguing the neighbourhood. The fact that she could have prevented the mugging by refusing to go out on that occasion does not itself diminish the wrongfulness of the mugging. The fact that J foresaw the mugging and could have easily avoided it does not vitiate the preventive or compensatory liability that the muggers bear for what they have done. So though J bears *causal* responsibility for being wronged in so far as she could

---

<sup>13</sup> This distinction belongs to Jeff McMahan, 'Liability, Proportionality, and the Number of Aggressors', *Ethics of War*, ed. S. Bazargan-Forward and S. Rickless (New York, [forthcoming]).

have avoided it, she is not morally responsible for the wrong of being mugged.

The same goes for P1 – though he is in a position to foresee being wronged, and though he can avoid this by refraining from committing A1, he is nonetheless not morally responsible for being wronged, provided that he genuinely did not want P2 to commit A2. Again, it is possible for P1 to be morally responsible for A2 in so far as it wrongs the victim of  $\phi$ , without thereby being morally responsible for A2 in so far as it wrongs P1.

We are now in a position to articulate the fundamental principle explaining why what P2 does is worse than what P1 does, all things being equal.

*The Principle of Wrongfully Imposed Accountability*

P1 has an alienable claim against P2 that P2 refrain from committing A2 on the grounds that (i) A2 makes P1 potentially liable to preventive or compensatory deprivations, (ii) P2 can avoid making P1 (or herself) potentially liable by refraining from committing A2, which she is obligated to do anyway since A2 is by hypothesis wrongful in so far as it causes  $\phi$ , and (iii) committing A2 in furtherance of P2's goals uses P1 by co-opting his actions in furtherance of a wrongful end which P1 repudiates (or is disposed to repudiate).

One might provide the following apparent counterexample to the principle of wrongfully imposed accountability:

'Firearm Dump'

An utterly reckless and lazy person (P1) wants to rid himself of his gun collection. Instead of selling them, giving them away, or properly disposing of them, he throws out each firearm from his car window as he drives through a crime-infested part of town. He does not want or intend anyone else to use these firearms; nor does she want to be made liable by such use. But predictably, one of the firearms is retrieved by a criminal (P2) who uses it to commit a murder.

In this example, P1 contributes to an agentially mediated wrong, i.e. the murder committed by P2. To see that P1 is morally liable to some preventive or compensatory deprivations, imagine that the only way to prevent the

murder is by imposing a substantial non-lethal harm on someone other than P2. Presumably the preventive harm we could permissibly impose on P1 is substantially greater than the preventive harm we could impose on some random bystander. (In the jargon of the literature on self-defence, the constraint of narrow proportionality, which says that the preventive harm imposed cannot be outsized relative to the harm prevented, is less stringent in its application to P1 than in its application to others.)<sup>14</sup> This indicates that P1 is morally liable to suffer preventive harm. So far, so good. But according to the principle of wrongfully imposed accountability, P2 wrongs P1 by making her accountable in this way. This might seem implausible in that it seems strange to say that P1 is among those whom P2 has wronged. If anything, we want to say that P1 has only herself to blame for being made liable to preventive and compensatory harms.

I believe we ought to say that P2 has indeed wronged P1 by acting in such a way as to make P1 potentially liable to preventive and compensatory deprivations. Our reaction that P1 ‘has no one to blame but herself’ is, despite initial appearances to the contrary, compatible with the view that P2 has wronged P1, provided we maintain that P1 is not entitled to complain to P2. P1 does not have the standing to express reactive attitudes of resentment and disapprobation towards P2, because doing so would be abjectly hypocritical. Even if P1 simply enjoys risk-taking and actually has no desire to end up responsible for the harms she enables, her conduct was so outrageously and pointlessly risky that for her subsequently to complain to P2 about being made liable would belie P1’s apparent flagrant disregard for the safety of others and *ipso facto* for her own liability. None of this is to say that P1 was not wronged. Rather, this is a case in which the wronged party is not among those positioned to complain about the wrongdoing. This analysis makes sense of our intuitions that P1 cannot complain to P2, without undermining the principle of wrongfully imposed liability.

If the principle of wrongfully imposed liability is correct, then *ceteris paribus*, in any case of agential mediation, P2 is potentially liable to greater preventive and compensatory harms than P1 because, by committing A2, she commits *two* wrongs – she wrongs P1 by making her potentially liable and

---

<sup>14</sup> The distinction between narrow and wide proportionality was first developed by McMahan. See McMahan, *Killing in War*, p. 20.

she wrongs the victims of  $\phi$ . This of course does not mean that P2 will always bear greater accountability than P1. Differences in intention, excusing-conditions, and risk-imposition between P1 and P2 can make P1 more accountable than P2 for  $\phi$ . But the fact that P2 is responsible not only for wronging the victims of  $\phi$  but also for wronging P1 means that, all things being equal, it is more likely that P2 bears greater accountability than P1 does.

To see this, suppose that the only way to prevent  $\phi$  is by breaking someone's wrist. But suppose that  $\phi$  is not itself so harmful as to justify breaking P2's wrist. To impose that preventive harm on P2 would, then, violate the constraint of narrow proportionality. But imposing that preventive harm on P2 achieves not only the good of averting wrongs to the victims of  $\phi$ , but also achieves the additional good of averting a wrong to P1. For if  $\phi$  is prevented, then P1 is no longer potentially liable. So there are actually two wrongs the prevention of which are included in the 'benefits' column of the proportionality calculation determining the permissibility of breaking P2's wrist – preventing the victim of  $\phi$  from being wronged, and preventing P1 from being wronged. And although the benefit of preventing the victim of  $\phi$  from being wronged will typically be much weightier than the benefit of preventing  $\phi$  from being wronged, the latter still carries significant weight, and can be enough to make P2 liable to a harm that would otherwise be disproportionate. The upshot is that P2's status as an intervening agent makes her liable to greater harms than those to which she would be liable if she had imposed the same risks but without thereby acting as an intervening agent.

### 3

My goal has been to show that there are principled reasons for thinking that what P2 does is worse than what P1 does. Whereas the literature on intervening agency focuses on whether P2's status as an intervening agent makes *P1's* conduct *less* bad, I turned this issue on its head by investigating whether P2's status as an intervening agent makes *P2's* conduct *worse*. And I argued that it does, by arguing that P2 wrongs not just the victims of  $\phi$  but P1 as well. On this view, it is not simply the fact that P2 is causally



‘closer’ to  $\phi$  that makes her actions more objectionable. Rather, it is what P2 does to P1 that makes her actions more objectionable.

This remains, of course, only a partial investigation of the moral relevance of intervening agency. The issue of whether P2’s status as an intervening agent makes P1’s conduct less bad relative to a case where P1 imposes the same risks absent an intervening agent remains unresolved. Neither have I discussed how intervening agency is related to *praise* (rather than blame) or how morally commendable (rather than objectionable) intervening agency affects accountability. These are issues for another time.<sup>15</sup>

[sbazargan@ucsd.edu](mailto:sbazargan@ucsd.edu)

---

<sup>15</sup> I thank Samuel Rickless for invaluable comments on an earlier version of this article. I presented an earlier version at the Rocky Mountain Ethics Congress, from which I received very helpful feedback. I am also grateful for helpful comments from an anonymous referee at *Utilitas*. I reserve my highest thanks for Craig Agule, who as my research assistant provided tremendous help on an earlier version of this article.