Political Philosophy

Renzo, Massimo. 2025. Why Manipulation is Wrong. *Political Philosophy*, 2(1): 176–203. DOI: https://doi.org/10.16995/pp.17395

Open Library of Humanities

Why Manipulation is Wrong

Massimo Renzo, Law, King's College London, UK, massimo.renzo@kcl.ac.uk

Philosophers have been paying increasingly more attention to the notion of manipulation. Perhaps surprisingly, however, most of the recent literature has focused on the conceptual question of identifying the features that make certain behaviours instances of manipulation, rather than on the moral question of why manipulation is wrong. Even more surprisingly, philosophers who have addressed the moral question have rejected the traditional view that manipulation is wrong because it undermines our efforts to exercise our practical agency and respond to our reasons for action. In this paper, I offer a defence of the traditional view. In doing so, I vindicate a further idea that is routinely invoked in discussions of manipulation but whose articulation has proved elusive: The idea that one reason manipulation is wrong is that it involves using people as means.

Political Philosophy is a peer-reviewed open access journal published by the Open Library of Humanities. © 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

OPEN ACCESS

Why Manipulation is Wrong

Massimo Renzo
Law, King's College London, UK

The notion of manipulation is increasingly invoked in relation to pressing issues discussed in the news and public discourse. Consider, for example, the recent debates on the use of propaganda and misinformation in democratic elections. Or the (more sober) discussions of whether governments are permitted to 'nudge' their citizens toward beneficial behaviour.[1] Or the recent 'Spy Cops' scandal, in which undercover British police officers induced unsuspecting individuals to form deceptive intimate relationships in order to infiltrate activist and political groups.[2] Running through these debates is the idea that certain attempts to influence conduct are objectionable in virtue of their being manipulative. Philosophers have thus been paying increasingly more attention to this notion, as reflected in the growing literature devoted to it. Perhaps surprisingly, however, this literature has focused primarily on the conceptual question of identifying the features that make certain behaviours instances of manipulation. Much less attention has been paid to the moral question of why manipulation is wrong.[3] Even more surprisingly, philosophers who have explicitly addressed the moral question have sometimes offered answers that, it's safe to say, significantly depart from how we normally think about the moral status of manipulation.

At least in philosophical circles, it's common to assume that the wrong of manipulation has something to do with the fact that manipulation somehow undermines our autonomy, or more broadly, our efforts to exercise our practical agency, understood as our capacity to respond to our reasons for action.[4] It's not uncommon, for example, to think that manipulation can undermine responsibility, and this thought is often driven by the assumption that being able to adequately respond to the reasons for action that apply to us is a condition for being responsible. Manipulation undermines

[1] Thaler and Sunstein 2021.

[2] Lewis and Evans 2010.

[3] For example, the recent *Stanford Encyclopedia* entry on 'The Ethics of Manipulation' lists 30 papers in its sections devoted to the conceptual question and only 5 in the (much shorter) section devoted to the moral question. On the distinction between the conceptual question and the moral one, see Noggle 2022 (though he uses different labels for it).

[4] Non-philosophers probably wouldn't put things quite in these terms, but I suspect that if asked, they would appeal roughly to the same idea to explain why manipulation is wrong.

responsibility by undermining this ability.[5] And yet, prominent attempts to answer the moral question reject this common assumption. In a classic paper, Sarah Buss has argued that the wrong of manipulation does not consist in the fact that manipulation compromises our autonomy or our efforts to exercise our capacity to respond to reason.[6] When we are manipulated, we might be wronged in several ways: we might be harmed, we might be prevented from having an accurate understanding of our circumstances, we might be placed in an unequal relationship with the manipulator. None of these wrongs, however, has anything to do with the fact that our response to our reasons for action is somehow compromised. For how would that even be possible? When we act as the manipulator directs us to, we are still responding to our own understanding of what we have reason to do, Buss argues. Of course, that understanding will be flawed insofar as it is based on the misleading evidence produced by the manipulator. But so what? Our understanding of what we have reason to do is often flawed, even when this is not the product of manipulation. We often rely on misleading evidence in choosing how to act and, because of that, we inevitably end up making choices that we would not have made had we possessed a better understanding of our situation. But if relying on misleading evidence does not compromise our practical agency when the evidence is the product of accidents or natural causes, why should it do so when it's the product of manipulation?

In a similar vein, Sophie Gibert has recently argued that the wrong of manipulation does not consist in the fact that manipulation compromises our efforts to adequately respond to our reasons for action. After all, it's often permissible to behave in ways that are intended to produce precisely this result. For example, businesses are allowed to employ good looking salespersons to induce us to purchase their products, despite the fact that the looks of a salesperson hardly constitute a good reason to buy a soda or a pair of shorts. And lawyers are allowed to choose charismatic witnesses and ask them questions that will paint them in a favourable light, even when these questions are not directly pertinent for the case. More importantly, Gibert argues, we can think of cases in which the way we respond to our reasons for action is *improved* by manipulation. Thus, far from being the key to understanding why manipulation is wrong, the idea that manipulation compromises our effort to respond to our reasons for action is not even a necessary condition of manipulation.[7]

---

[5] Things are in fact more complicated than that (Renzo 2023).

[6] Buss 2005.

[7] Gibert 2023.

My aim here is to defend the traditional answer to the question of why manipulation is wrong. Indeed, I will defend an especially strong version of this answer. I will argue that there is a distinctive wrong of manipulation, which consists in producing a particular type of interference with our efforts to exercise our practical agency and adequately respond to our reasons for action. Moreover, I will vindicate an idea that is routinely invoked in discussions of manipulation but whose articulation has proved elusive: The idea that one reason manipulation is wrong is that it involves using people as means. To be clear, I don't think these are the only wrong-making features of manipulation. There are many reasons why manipulation is wrong. But the reasons discussed here deserve a central place, I will argue, in any plausible answer to the moral question of why manipulation —at least manipulation of a certain kind— is wrong.

Although the aim of this paper is to answer the moral question, my discussion will not be confined to it. This is for two reasons. One is that any account of why manipulation is wrong will have to rely (how could it not?) on some understanding of what it takes for a given behaviour to count as manipulative.[8] The second reason is that the distinctive wrong of manipulation is, as I will argue, grounded in a feature of manipulation that has been ignored so far in discussions of the conceptual question. If I'm right about this, at the end of the paper we will have learnt something important not only about what makes manipulation wrong but also about what manipulation is.

I'll proceed as follows. I start by outlining four different ways in which manipulation can be said to interfere with our efforts to exercise our practical agency (section I).[9] This is important because these types of interference are structurally very different and any attempt to provide an account of the wrong of manipulation that is oblivious to these differences is likely to be off target. I then focus on one of these types of manipulation, which tends to be regarded as paradigmatic in the literature. I argue that providing an account of the wrong associated with this kind of manipulation requires paying attention to a particular way in which our interest in exercising our practical agency can be undermined (sections II and III). Finally, I offer an account of why manipulation violates the means principle (section IV).

---

[8] It's less clear that the reverse also holds. Perhaps it does if manipulation is always *pro tanto* wrong. If so, we should at least rule out as implausible any accounts of the nature of manipulation that leave open the question of whether manipulation is *pro tanto* wrong. Philosophers disagree, however, as to whether manipulation is always *pro tanto* wrong (Baron 2014; Wood 2014; Noggle 2022).

[9] I am grateful to an anonymous reviewer for helpful comments that prompted me to include this section.

I. VARIETIES OF MANIPULATION

The first thing to say in answering the moral question is that there cannot be a single answer to it. This is because manipulation is a complex phenomenon and can take different forms. The cases that come to mind first tend to be those involving deception, as when Iago places Desdemona's handkerchief in Cassio's room to get Othello to doubt her fidelity. But many instances of manipulation do not involve deception. For example, if my aim is to induce you to refuse a vaccine that has been proved to be safe, one option is targeting your beliefs, say, by misrepresenting the data on the safety of the vaccine; another is intervening on your feelings, say, by constantly bringing up some modest side effects of the vaccine so as to make you nervous about it, despite your knowledge that the vaccine is perfectly safe.[10] Or think of someone who's browbeaten into deferring to their spouse's needs, or someone who's brainwashed into joining a cult by a leader that employs sophisticated forms of psychological punishment and reward. These behaviours seem manipulative in virtue of the fact that they involve forms of undue psychological pressure, not in virtue of the fact that they are deceptive.[11]

And there seem to be forms of manipulation that involve neither deception nor psychological pressure, but what is sometimes called 'arational influence', as when advertisers use particular fonts or graphics in an effort to nudge the audience toward certain behaviours, or when a pharmaceutical company decides to produce tranquillizers with a certain pill colour because of a subconscious tendency to associate that colour with tranquillity.[12]

Presented with this range of cases, a natural response would be to try and explain the wrong of manipulation by appealing to the wrong of deception, psychological pressure or arational influence, or perhaps a combination of these wrongs. This would be relatively straightforward. For example, we already have various accounts of why deception is wrong. We could pick our favourite one and use it to explain why *deceptive manipulation* is wrong. Then do the same with psychological pressure and arational influence.

This strategy, however, is ultimately not very promising. To see why, consider the case in which I make you nervous about getting a vaccine by constantly bringing up its modest side effects. I've introduced this as a case of manipulation involving

---

[10] Cave 2007.

[11] Of course, we can think of behaviours that are both deceptive and involve the exercise of psychological pressure.

[12] Cohen 2018.

psychological pressure, and this seems correct if we assume that it's the mere repetition of the side effects that wears you down, making you fall prey to what you know is an irrational fear. But could this not be an instance of arational influence? After all, your belief is not rationally supported by what I say, just like your choice to drink a particular brand of soda is not supported by the enticing images used by advertisers. And are we really sure that it's not an instance of deception? There are certainly versions of the case that can be aptly described this way. For example, if my repetitive behaviour was effective by somehow generating in you the belief that the vaccine is more dangerous than the available data suggests (say, because you take my insistence to be evidence of the seriousness of that danger), this would seem an instance of deceptive manipulation after all.

How should we go about deciding which of these categories applies? We need to look at the specific way in which my interference affects your attempt to respond to your reasons for action. Does it succeed by wearing you down so as to make you feel nervous, despite the fact that you know the vaccine is safe? By circumventing your capacity to respond to reasons altogether through the exploitation of some psychological bias that is not reason-responsive? By generating new reasons for belief? Once we have answered these questions, however, it's unclear what is to be gained by invoking the categories of psychological pressure, arational influence or deception. Indeed, the worry is that doing so can be a source of confusion. The notion of psychological pressure, for example, is used to refer to cases that are structurally very different. The sort of interference produced when a wife is browbeaten into deferring to their spouse's needs is importantly different from the one produced when a student is pressured into having sex with her teacher in exchange for higher grades or the one produced when someone is brainwashed into joining a cult.[13] That they can be all referred to as instance of psychological pressure risks obfuscating this fact.

This is especially important in the context of determining why these behaviours are wrong, since we might be tempted to think that the wrong perpetrated in all these cases is the same insofar as they are all instances of 'psychological pressure'. That clearly would be a mistake. The difference between these behaviours is not simply in the degree of pressure imposed, but in the type of interference they produce. Appreciating this difference is essential to understanding why these behaviours are wrong.

Let me then suggest a different approach. Instead of starting with the categories of deceptive manipulation, psychological pressure and arational influence, let's start

[13] For a discussion of the first two cases, as well as other cases of manipulation that take the form of psychological pressure, see Baron 2003.

by distinguishing four types of manipulation based on four different strategies a manipulator (M) can employ to interfere with his victim (V)'s efforts to respond to her reasons for action:

Type 1: *Cases where M shapes V's circumstances so as to guide V to φ by creating new actual reasons for V to φ.*

Exploitative offers operate this way. The teacher offering higher grades in exchange for sex is creating new reasons for the student to sleep with him. These are actual reasons that the student didn't have before the incentive was offered. Coercion also operates this way. By threatening to shoot your child, the robber creates new reasons for you to open the safe; by placing a hospital next to the munition factory, the insurgents create new reasons for the army not to shoot the munition factory.

Type 2: *Cases where M shapes V's perception of the circumstances so as to get V to φ because φ-ing will seem to V what they have reason to do in light of their values and goals. However, the perception of the circumstances induced by M is misleading.*

Cases of deception, like Othello's, have this structure. But deceiving someone involves deliberately causing them to believe something that is not true, and this is not the only way this type of manipulation operates. For example, instead of intervening on your beliefs, I might intervene on your feelings, as we've seen in the case in which I make you nervous about a vaccine you know to be safe by constantly bringing up its side effects.

The difference between Type-1 and Type-2 manipulation is that while the former attempts to influence behaviour by creating *actual* reasons for the manipulee, the latter does so by creating *apparent* reasons, i.e. reasons that we understandably take ourselves to have in light of mental states that, given the interference, we are justified in having.[14] If the vaccine's side effects are modest, we don't have reason to be too worried about them. Manipulators cannot change that by presenting us with misleading evidence or by playing with our emotions. At most, they can create apparent reasons for scepticism.

Type 3: *Cases where M shapes V's ideals and goals so as to get V to φ because φ-ing is what someone with those ideals and goals has reason to do. However, V's ideals and goals are defective in that the process through which V has adopted them somehow undermines their normative force.*

[14] On the notion of apparent reasons, see Parfit 2011.

The difference between this type of manipulation and the previous one is that here it's the manipulee's ideals and goals that are being targeted. Type-2 manipulation, by contrast, targets the manipulee's perception of what their own ideals and goals require. In Type-3 cases, the manipulator instils in the manipulee new values or cares; in Type-2 cases, the manipulator takes the values and cares of the manipulee for what they are and confuses her about what they call for.

Similarly, the difference between Type-1 and Type-3 manipulation is that the former takes the ideals and goals of its target for what they are and triggers reasons that, given those ideals and goals, will call for a certain response. (Knowing that my son wants to join my law firm, I make clear that a position will become available if he chooses to live with me rather than with his mother.) Type-3 manipulation, by contrast, implants new ideals and goals which, once adopted, generate new actual reasons for action. (Knowing that my son is inclined to become a musician, I inculcate in him a desire to be a lawyer instead by playing on the fact that all the members of our happy family have been lawyers for generations.)

Most cases of brainwashing and indoctrination are (extreme) examples of Type-3 manipulation, as are the cases typically discussed in the free-will literature.[15]

Type 4: *Cases where M gets V to φ by triggering in V psychological processes that are not reason-responsive.*

This group includes a range of diverse techniques that exploit irrationalities we are prone to. Examples include some of the cognitive biases discussed in behavioural economics and social psychology, such as framing effects (as when a doctor frames the risk of surgery in terms of survival rates rather than mortality rates), priming bias (as when the polling station to vote for a proposition about increasing school funding is placed in a school) and anchoring effects (as when a negotiator makes a low opening offer in an attempt to get the negotiation to gravitate toward that amount).[16] But in some cases, the irrational association can be simpler and more direct, as in the case of the pills manufacturer, or in cases of advertisements that create a connection between a product and certain ideals of success or happiness.

The distinctive feature of this type of manipulation is that, unlike the previous three, it does not operate via the creation of normative reasons. A patient does not acquire any new reasons for action, actual or apparent, in virtue of the fact that the risk of surgery

[15] Fischer 2004; Pereboom 2014; Mele 2019.

[16] Kahneman 2011. See also Doris 2002.

is framed in terms of survival rates rather than mortality rates.[17] Nor do consumers acquire reasons to drink Coca Cola in virtue of the fact that Santa is pictured as drinking Coca Cola on a billboard.

As far as I can tell, these cases are what philosophers have in mind when they talk about arational influence.

To be clear, this classification is not meant to provide an account, not even the outline of an account, of the types of manipulation falling in each of these categories. It is only meant to distinguish four different ways in which we can interfere with someone's efforts to exercise their practical agency when we manipulate them.[18] More needs to be said about each of these types of interference to explain the sense in which they are manipulative. For obviously I can create new reasons for action, actual or apparent, for others without manipulating them. I can create actual reasons for you to pass me the salt simply by asking. I can create (merely) apparent reasons for you to believe that I'm home by forgetting to shut the front door. As for cases in which we implant ideals and goals in others non-manipulatively, it is often remarked that various processes of education and socialization that we undergo as children involve precisely that. Certain values and cares are instilled in us by our parents and society through methods that, generally, are regarded as different from brainwashing or indoctrination.[19] Finally, there are plenty of cases in which our cognitive biases are activated regardless of other people's intervention, and those clearly would not be cases of manipulation, if only because manipulation requires a manipulator.

Distinguishing the four types of interference with our practical agency identified above is only the first step in providing an account of manipulation. For each type, we need an account of what it takes for the interference in question to count as manipulative. In the next two sections, I will provide an account of this sort with respect to Type-2 manipulation, i.e. cases in which someone's perception of their circumstance is

[17] Unless the framing in itself is meant to be evidence of something else (Levy 2019).

[18] Once these four types of manipulation are distinguished, we can see how some of the disagreements in the literature about issues such as whether manipulation and coercion are qualitatively different (Baron 2014; Wood 2014), or whether manipulation bypasses our capacity to respond to reasons (Buss 2005; Cave 2007; Gorin 2014a), are simply the result of philosophers talking past each other. The claim that coercion and manipulation are on a continuum, rather than qualitatively different, is plausible with respect to Type-1, but not with respect to other types of manipulation. Similarly, the claim that manipulating someone involves bypassing their capacity to respond to reasons is plausible with respect to Type-4; less so with respect to other types (though much will depend on how the claim is interpreted).

[19] Though explaining the difference is far from easy (Haji and Cuypers 2008).

interfered with so as to get them to φ because φ-ing will seem to them what they have reasons to do. (In talking about 'manipulation', I will always refer to this type of case.) I will then build on that account to explain why this type of manipulation is wrong. Since many classic cases of manipulation, including Othello's, take this form, this seems a like a good place to start in answering the moral question. Providing accounts for the other types of manipulation is a task for another day.

II. MANIPULATION AND PRACTICAL AGENCY

If our aim is to vindicate the traditional idea that the wrong of manipulation consists in undermining our efforts to exercise our practical agency (understood as the capacity to adequately respond to our reasons for action)[20] our first order of business should be to spell out two ideas: what it takes to exercise our practical agency well, and how our efforts to do so are affected by manipulation. Start then with the idea that the point of exercising our practical agency is to adequately respond to the reasons for action that apply to us.[21] Some of these reasons (say, the reason to study for the bar exam) depend on choices and commitments we have made; others (for example, the reason not to inflict unnecessary harm) do not.[22] But regardless of their origin, at any point in our life, we are confronted with a number of reasons for action that apply to us. Exercising our practical agency is a matter of adequately responding to such reasons.[23] What does that involve?

One view is that adequately responding to our reasons for action is a matter of appropriately responding to our apparent reasons, i.e. what we justifiably believe our reasons to be, given our evidence (or, perhaps, given our justified beliefs, or given what we know) at the time of the action. This is the sort of view that Buss seems to have in mind when she writes that 'whether an instance of practical reasoning is self-determined is a matter of whether it is really the agent herself who is doing the reasoning. And this would seem to depend on whether she determines her response to the considerations that figure in her reasoning—not on how the considerations to which she responds relate to reality.'[24]

---

[20] Reason-based accounts of manipulation are also offered by: Mills 1995; Yaffe 2003; Gorin 2014b.

[21] Scanlon 1998; Parfit 2011; Raz 2000; Raz 2011.

[22] Raz 1986, pp. 385–90; Chang 2013.

[23] At least to the extent that we are capable of responding to reasons. There is no clear sense in which reasons for action apply to infants or, say, people in a coma.

[24] Buss 2005, p. 214.

If we take this view, it's a short step to the conclusion that manipulation cannot undermine our efforts to exercise our practical agency and respond to our reasons for action. Even if you are responsible for manipulating me into believing that snake oil is a great remedy for rheumatism, I will only buy the vial you're dangling in front of me if I decide that doing so makes sense in light of my own understanding of the evidence I'm presented with. In that respect, the way in which I respond to my reasons for action when I decide to buy snake oil is no more defective when it is prompted by your manipulation than when it is entirely self-induced, the product of my poor understanding of the chemical properties of snake oil. It's in this sense that, as Buss puts it, when the manipulee acts, 'she will act for her own reasons if she is to act at all'.[25] If exercising our practical agency is simply a matter of responding to our apparent reasons, the question of whether those reasons track our actual reasons (the reasons we in fact have) is beside the point in determining whether we are managing to respond to our reasons for action.

But is it true that exercising our practical agency is simply a matter of responding to our apparent reasons? How we respond to our apparent reasons is certainly important in a number of ways. By showing that our conduct is well supported by a sound understanding of our apparent reasons for action we show that it was rational for us to behave as we did.[26] And, at least in some cases, this will also mean that we are not blameworthy if our conduct turns out to be morally wrong. Insofar as it was rational for us to behave as we did, our conduct does not manifest a bad character or a morally flawed attitude.[27] But in exercising our practical agency, our aim is not only to avoid irrationality or blame. *Qua* practical agents, our primary concern is doing what we have actual reason to do, rather than merely being rational or being blameless.[28]

---

[25] Ibid., p. 213.

[26] Parfit 2011, ch. 5; Broome 2021. This kind of justification drives most of Buss' discussion, as I understand it. Notice that while Buss often frames her view in terms of whether manipulation undermines autonomy, she identifies being autonomous with being a rational agent (what I call a 'practical agent') who gets to respond to their own reasons for action. For example, she defends her view by writing that 'if … rational agents can be said to "fully determine their actions according to reasons" whenever they take their actions to be adequately justified, then deceit and manipulation cannot prevent someone from being a truly autonomous, truly rational agent' (Buss 2005, p. 213). I intentionally avoid discussing autonomy here, since most of the contemporary discussion understands manipulation as interfering with our capacity to respond to reasons, rather than our autonomy. But I believe the view I outline could be reframed in terms of autonomy.

[27] Gardner 2007; Scanlon 1998, pp. 279–80.

[28] Dworkin 2011; Raz 2011; Wall 2021.

Consider:

> *Coffee*: Because of a mix up in a distribution warehouse, the bag labelled 'Sugar' sitting in Amina's kitchen contains cyanide instead. Karim comes over for coffee and Amina puts some cyanide into his cup.[29]

Pouring some of the content of the sugar bag into Karim's coffee is perfectly rational for Amina, and she is not blameworthy for doing so. But clearly there is a sense here in which she's failing to adequately respond to her reasons for action when she does that. After all, Amina's reasons not to act as she does persist despite her excusable ignorance. But if so, in acting against these reasons, Amina fails to respond to them. For how could those reasons have normative force if acting against what they call for does not count as a failure, even if an excusable one, to respond to them?

One might be tempted to reply that the normative force of those reasons consists in the fact that Amina ought to be moved by them (absent countervailing reasons) *if she were aware of them*. But this won't do. At best, this reply would explain the force of those reasons for someone who is aware of them; not for Amina, who —we're stipulating here— is not. In other words, the reasons Amina has not to pour the content of the sugar bag into Karim's coffee are not plausibly understood as requiring the following: '*If Amina were aware of the fact that she has reasons to refrain from pouring the content of the sugar bag into Karim's coffee, she ought to do so.*' That would make Amina's reasons not to pour the content of the sugar bag conditional. Amina's reasons, however, are not conditional in this way. They call for her to refrain from pouring the content of the bag into Karim's coffee even while she is unaware of them.[30]

So, *contra* Buss, there is a clear sense in which we fail to adequately respond to the reasons for action that apply to us when we fail to do what we have actual reason to do, even if our conduct is justified as a way of responding to our apparent reasons. This, however, is not enough to vindicate the traditional account of the wrong of manipulation if, as Gibert argues, there are cases in which the way we respond to our reasons for action is improved, rather than compromised, when we're manipulated.

This is the sort of case Gibert has in mind:

> India and Sebastian are having a serious discussion about who will make career sacrifices next year to care for their kids. Sebastian has recently observed that India tends

[29] I borrow this example from Harman 2011.

[30] Renzo 2025, p. 186.

> to undervalue her own preferences and needs, but that she is slightly less inclined to do so when she is angry at someone. Before their conversation, Sebastian ensures that India overhears him on the phone with his mother, divulging some of their marital issues that they have agreed to keep private. As a result, India gets angry and gives more appropriate weight to her preferences in their subsequent discussion.[31]

Here, Gibert argues, Sebastian's manipulative interference does not seem to have the effect of compromising India's response to her reasons for action. Quite the opposite, his intervention is precisely what enables India to appropriately respond to them. If so, compromising our efforts to appropriately respond to our reasons for action is not even a necessary condition of manipulation, let alone the central notion on which an account of the wrong of manipulation ought to be grounded.

But is it true that Sebastian's intervention has enabled India to appropriately respond to her reasons for action? In one respect it has, since it has led her to insist on the distribution of responsibilities that she has reason to seek out. In another respect, however, it hasn't. Being angry about someone's indiscretion is not a good reason to value one's own needs more, and thus inducing India to do so as a response to those feelings of anger is a way of inducing a flawed response to her reasons for action. What should move India to value her own needs more is rather her recognition that those needs matter no less than Sebastian's. So, while there is an important lesson to be learned from this case, it's not the one that Gibert suggests. The case does not provide a counterexample to the traditional view that the wrong of manipulation is grounded in its compromising our efforts to respond to our reasons for action. Rather, it draws our attention to a feature of the traditional account that is often neglected by its own defenders. The account should be interpreted as concerned not only with the question of whether the victims of manipulation end up acting as they have reason to, but also with the question of whether in acting as they have reason to, they are motivated by the right reasons.[32]

Now, the value of being motivated by the right reasons has received significant attention in philosophical discussion. That value, however, is normally located in the fact that being motivated by the wrong reasons deprives our conduct of moral worth.[33] What Gibert's case shows is that even when the moral worth of our conduct is not

---

[31] Gibert 2023, p. 359.

[32] I develop this criticism in Renzo Forthcoming.

[33] Arpaly 2002; Markovits 2010.

at stake, there is an important sense in which we can fall short in responding to our reasons for action when we are motivated in the wrong way.[34]

Here's another example. Mozart had reasons to express his creativity by composing music. Now imagine that instead of being moved by these reasons, Mozart's only motivation in composing his beautiful music was that this is what his father Leopold expected of him. This is not a moral failure, insofar as meeting one's parent's expectations is an acceptable reason to compose music. Still, there seems to be a clear sense in which this Mozart falls short in responding to the reasons that apply to him. Plausibly, the reasons we have to express our creativity should take precedence in motivating us to produce art over the reasons we might have to fulfil others' expectations. If so, the way in which we respond to our reasons for action is flawed when we are moved by the latter rather than by the former. What kind of failure is that? The best way to put it is, I think, to say that there is a sense in which my imaginary Mozart is alienated from his own reasons for action. Although his conduct conforms to what those reasons call for, he fails to grasp how he is to be moved by them. His reasons make certain demands on him, and he fails to recognize those demands for what they are, even if his conduct is not at odds with them. There is thus, a disconnection between how the reasons that apply to him are meant to govern the exercise of his practical agency and how he in fact exercises it.[35]

If I am right, we should conclude that there are two ways in which we can fall short in responding to our reasons for action: We can fail to do what we have (actual) reason to do or we can do what we have reason to do for the wrong reason.[36] Falling short in

---

[34] One might be tempted to resist this point by pointing out that in the case at hand the moral worth of India's conduct *is* at stake, since we can plausibly imagine that her failing to give sufficient weight to her own needs reflects a subservient attitude toward Sebastian. Notice, however, that this is not part of the case as described by Gibert, and while we can imagine versions of the case in which India's behaviour is indeed explained by a lack of self-respect, we can also imagine versions that feature other explanations. I offer one in Renzo Forthcoming.

[35] I defend this view of practical agency in Renzo 2025. Notice that the notion of alienation I'm using here does not refer to a psychological phenomenon, such as feeling passive with respect to certain mental states we experience as extraneous to us. My imaginary Mozart might not experience any of that. He is alienated from his reasons for action in the sense that there is a disconnection between how those reasons are meant to govern the exercise of his practical agency and how he in fact exercises it. On this 'objective' sense of the notion of alienation, see Hardimon 1994, pp 119–22.

[36] Doing what we have reason to do does not mean responding to *all* the reasons that apply to us. Reasons for action often conflict, in which case we need to trade-off between them. And some reasons are pre-empted in our deliberation in virtue of certain features of practical reasoning (for example, when we have promised something). But at the end of the day, there is a set of undefeated reasons for action that emerges as the one we ought to respond to, all things considered (Scanlon 1998; Raz 2011).

either way involves a sort of alienation from our reasons for action. We are alienated from them insofar as we fail to recognize them for what they are, and thus fail to be moved by them in the way they call for.[37]

Armed with this understanding of what it takes to exercise our practical agency well, we can now offer the following idea as a first pass of an account of manipulation: Manipulating someone involves intentionally causing them to suffer the sort of alienation from their reasons for action described above in order to get them to act in certain ways. So, I might manipulate you:

a) by intentionally getting you to fail to do what you have (all-things-considered)[38] reason to do – say, by inducing you not to get a vaccine when in fact you have reason to get one; or

b) by intentionally getting you to do what you have reason to do for the wrong reasons – say, by misleading you into believing that being vaccinated is a condition of working at your new job.[39]

---

[37] An interesting difference between India's case and Mozart's is worth noting here. While Mozart's alienation can be easily located in the fact that fulfilling his father's expectation is not the reason that should move him to compose music, things are more complicated in India's case. After all, the fact that she values her needs more *is* the reason that should move India to insist on a more equal distribution of domestic responsibilities. The problem here is that India values her needs more because of her anger, rather than because she appreciates the fact that her needs matter no less than Sebastian's. This is why she fails to be moved by the reasons that apply to her in the right way. As Julia Markovitz notices, 'we generally act for chains of dependent motivating reasons, running from the less to the more fundamental' (Markovitz 2010, p. 227). For example, although my running into the burning house is motivated by my intention to save the child trapped in it, this is not the end of the story if we're interested in what motivates my conduct. We still need to know what my motivating reason for saving the child is. Is it that I care about his life or that I want to look good with the neighbours? And if it is that I care about his life, what's the motivating reason for that? Is it that I value his life in itself or that I stand to benefit from it somehow (say, because I'm a well-paid trustee)? The point here is that defective reasons can appear at different levels along the motivational chain that leads us to act. The faulty link in the motivational chain that leads India to act is located at a more fundamental level than the one in Mozart's motivational chain. An interesting question, which I cannot consider here, is how this feature affects the degree to which we fall short in responding to our reasons for action.

[38] Henceforth, I will take this qualification for granted.

[39] I originally sketched this view in Renzo 2021 and discuss (b) at length in Renzo Forthcoming. The importance of (b) is recognized by Moti Gorin (2014, 2018). Notice, however, that Gorin understands (b) in subjectivist terms. For him, I manipulate you when I ensure that your apparent reasons are misaligned with what *I believe* your actual reasons are. By contrast, my understanding of (b) is objectivist: I manipulate you when I intentionally ensure that your apparent reasons are misaligned with what your actual reasons in fact are.

Importantly, our reasons for action are typically mediated by the reasons we have to form certain beliefs, emotions and attitudes. Normally, we form the intention to φ because we believe that φ-ing is what we have reason to do, or because φ-ing is called for in light of certain attitudes and emotions of ours. Manipulation involves interfering with our effort to respond to our reasons for action by compromising the way in which we respond to our reasons for beliefs, attitudes and emotions. It's in this sense that manipulation involves 'worsening someone's deliberation',[40] or inducing them to 'make a mistake',[41] to 'fall short of certain ideals that regulate their behaviour',[42] or to 'fail to track the reasons for action that apply to them'.[43]

III. WHY MANIPULATION IS WRONG

Inevitably, we've had to paint with a broad brush. Hopefully, however, the main elements of the picture will be sufficiently clear for our purposes here. We now have a sense of why exercising our practical agency is not exclusively a matter of responding to our apparent reasons for action. There is an important way in which we fall short, *qua* practical agents, when we fail to adequately respond to the actual reasons for action that apply to us, either because we fail to do what we have reason to do, or because we do what we have reason to do for the wrong reason. And we have a sense of how manipulation undermines our effort to exercise our practical agency in this way. At first approximation, we can say that manipulating someone involves intentionally[44] interfering with their efforts to adequately respond to their reasons for action in order to induce them to fall short in one of these two ways.[45]

---

[40] Hanna 2015.

[41] Noggle 2020.

[42] Noggle 1996.

[43] Gorin 2014b. There are significant differences between these views, which need not concern us here. Hanna's is the closest to the picture I've outlined, since it's the only one for which whether we are manipulated depends on whether we are induced to act in ways that are at odds with our actual reasons for action, rather than the reasons for action that the manipulator believes we have.

[44] Or perhaps recklessly (Baron 2003). Henceforth, I will take this qualification for granted.

[45] I focus here on reasons for action because typically manipulators interfere with our effort to respond to our reasons to form beliefs, attitudes and emotions with an eye to affecting how we act. But the view could be easily adapted to account for cases in which the only aim of the manipulator is to affect the beliefs, attitudes or emotions of his target, where this is not part of an attempt to induce the target to act in a certain way.

So, now we might think that an account of the wrong of manipulation is almost in hand. If exercising our practical agency is a matter of responding to the actual reasons for action that apply to us, anyone who intentionally prevents us from doing so undermines our efforts to exercise our practical agency. Isn't this, plain and simple, the reason why manipulation is wrong? Perhaps that's why philosophers have not felt the need to spend much time discussing the moral question. Once we give the sort of answer outlined above to the conceptual question, we might think that the answer to the moral question is more or less obvious.

This would be too quick, though. For even if we assume that we have an interest in exercising our practical agency well and adequately respond to the actual reasons for action that apply to us, it's implausible to think that we are the victims of wrongful manipulation —even *pro tanto* wrongful manipulation— whenever this interest is set back. This becomes clear once we consider that so many of the influences we are constantly exposed to in our interactions with others have precisely this effect on us, whether it's intended or not. Indeed, as Buss reminds us, the same effect is produced all the time by natural causes as well.[46]

Now, one thing to say in reply to this sort of worry is that the reason we are not wronged when our efforts to respond to our actual reasons for action are frustrated by natural causes (or by interactions with others that are merely accidental) is simply that in order for a wrong to be present, someone needs to act with a given mental state. To see this, consider for a moment the way in which we're wronged by harmful conduct. If you smash my beautiful bay window with a rock, you wrong me, provided that you acted intentionally (or at least recklessly or negligently). Our confidence in this judgment is not —and should not be— shaken by the thought that if the same rock had been carried by the wind, I would not be wronged. It would be a mistake to argue that insofar as the same harm could also have been produced by natural causes, and in that case I would not have been wronged by it, then the harm must be playing no role in explaining the wrong I suffer when the harm is intentionally (or recklessly) produced by you.

The same can be said for the way in which, say, misleading evidence undermines our efforts to respond to our reasons for action. We can imagine a version of *Othello* in which Desdemona's handkerchief is blown by the wind in Cassio's room, and this leads Othello to doubt her fidelity. In this version of the story, Othello is not wronged even if, as a direct consequence of this unlucky event, he fails to exercise his practical agency and adequately respond to what he has reason to do in the same way he does

[46] On this problem, see also Yaffe 2003.

in Shakespeare's story. Still, this should not lead us to doubt that this failure does play a role in determining why Othello is wronged in Shakespeare's story (where the handkerchief is intentionally planted in the room by Iago). It would be a mistake to argue that insofar as the same failure could have been produced by the wind, and in that case Othello would not have been wronged, then that failure must be playing no role in explaining why Othello is wronged in the version of the story where the failure is intentionally produced by Iago.

This reply only goes so far, however. It helps vindicating the idea that undermining our efforts to respond to our reasons for action is a wrong-making feature of manipulation, but it fails to vindicate the idea that there is a distinctive wrong we suffer when we are manipulated. This is because I can wrongfully undermine your effort to respond to your reasons for action in ways that are not manipulative. For example, I could intoxicate you or, like in the movies, stun you by hitting you in the head with a bottle. I take it that even if doing so leads you to make a mistake, to fall short of the ideals that regulate your behaviour, or to fail to track the reasons for action that apply to you, this would not be an instance of wrongful manipulation. It would not be an instance of wrongful manipulation because it wouldn't be an instance of manipulation to begin with.

To be an instance of manipulation, our interference with someone's attempt to exercise their practical agency and adequately respond to their reasons for action must take a particular form. Manipulation involves exploiting someone's deliberative faculties in order to produce the relevant misalignment between the reasons that ought to move them —their actual reasons for action— and the reasons that do move them. When I manipulate you, *you* are the one who comes to decide that you should not get the vaccine, though you're led to make this decision by the misleading evidence I have deliberately supplied you with. Similarly, Tom Sawyer's friends decide themselves to spend the afternoon painting Aunt Polly's fence, though Tom is the one who expertly leads them to make that decision by feigning enthusiasm for the supposed privilege of performing that task.

One reason the wrong of manipulation is so insidious is precisely that it is by exploiting our own deliberative faculties that manipulators undermine how we exercise our practical agency. In exercising our practical agency, our aim is to do what we have reason to do, and to do so for the right reasons. Manipulators mislead us about what those reasons are by ensuring that our apparent reasons for action are misaligned with our actual ones. As a result of that, our effort to exercise our practical agency well becomes self-defeating: the more we try to appropriately respond to our reasons for action, the more we fail to do so.

This is the feature that is overlooked in existing accounts of the nature of manipulation — those which understand manipulation in terms of certain end states, such as falling short of certain ideals, or failing to track certain reasons for action. And this feature, I want now to suggest, is the key to answering the moral question. The distinctive wrong we suffer when we are manipulated is not merely that we are led to make a mistake in our deliberation, so that we fall short in how we respond to the reason for action that apply to us. It's that our own capacity to exercise our practical agency is being exploited by the manipulator to produce this result.

To see this, compare Othello's fate with the one he would have suffered if, instead of manipulating him, Iago had hypnotized him and ordered him to kill Desdemona. In that case, although there is a sense in which Iago would have exploited Othello's agency, he would not have exploited Othello's capacity to respond to the reasons for action that apply to him. But one way in which Othello's fate is so tragic is precisely that he is the one who makes every decision that leads him to his ruin. What makes manipulation especially objectionable —and what is distinctive of this kind of interference— is that the capacity of the manipulee to respond to their reasons for action is co-opted by the manipulator in this way. It's the manipulee's own efforts to exercise their practical agency well and respond to their own reasons that ultimately condemns them to fail to do so, either because they end up acting against what they have reason to do or because they end up doing what they have reason to do for the wrong reasons. Their practical agency is thus not simply damaged or impaired but turned against itself. It's turned against itself insofar as their failure to adequately respond to the reasons for action that apply to them arises from their very own effort to exercise their practical agency well so as to successfully respond to those reasons.[47]

Victims of manipulation need not be aware of this fact, of course. Indeed, typically they are not, since manipulation often takes the form of covert influence.[48] In these cases, the victim responds to the apparent reasons created by the manipulator because, like Othello, they mistakenly believe that those reasons do track their actual ones. However, there can also be cases in which the victim's role in the process of their own manipulation becomes clear to them. For example, I might realize that a political candidate constantly talks about immigrants in order to make me nervous about them, and yet find, to my surprise, that his strategy is working. It's working even if I know that countries like mine in fact benefit from immigration. Here, when I get nervous

[47] A similar idea is hinted at in Coons and Weber 2014, p. 16, fn. 11. I'm grateful to a reviewer for bringing this to my attention.

[48] Goodin 1980; Susser, Roessler, and Nissenbaum 2019.

about immigrants, my conduct is the product of my own (inapt) response to what I know have reason to feel. I know that the apparent reasons to be nervous the candidate is producing with their behaviour do not track actual ones, and yet I end up treating them as if they did.[49]

Now, a significant feature of this case is that in addition to *being* alienated from my own reasons for action, I also *experience* this alienation. Not only is my practical agency turned against itself; I am also conscious of this process as I contribute to it. This, however, should not lead us to revise the account offered so far. For in this case too, the central wrong I suffer is not that I have been made to feel in a certain way, but rather that my practical agency is exploited to undermine my effort to respond to what I have reason to feel and to do.

To see this point, notice how my feeling nervous is not something that simply happens to me, like my feeling hungry or feeling cold. If I feel hungry, being aware that I just had a big meal gives me no reason to revise that feeling.[50] Feeling hungry is a purely empirical phenomenon I experience, regardless of whether I think I have reason to do so. My capacity to respond to normative reasons has no role to play here. Feeling nervous about a threat is different. Being aware that immigrants are not a threat does give me reason to revise that feeling. Exercising my practical agency well involves correctly responding to this fact, modulating my emotional responses according to my own assessment of what I have reason to feel.[51] When the candidate's interference prevents me from doing that, it affects me not by making me experience a brute emotion of panic, but rather by compromising how I respond to my own assessment of the situation. It's in this sense that my practical agency is coopted by the manipulator and turned against itself. Being nervous is a flawed response to my normative reasons that is produced by my own effort to exercise my practical agency well, where this effort has been thwarted by the manipulator's intervention. Since I am aware of this process, I will likely experience a further level of estrangement from my practical agency —one

[49] Cases like these get close to those of arational influence, though I would want to preserve the distinction between the two. Even if getting nervous about immigrants is not a rational response to anti-immigration propaganda when we know that immigrants are not a threat, the response does not seem to be on a par with central cases of arational influence (such as those involving framing bias), in which the manipulator does not even pretend to offer genuine reasons. Of course, this point cannot be adjudicated without an account of arational influence. But even if I'm wrong about this, that wouldn't cast doubts on my account. Rather, it would show that some forms of arational influence are instances of Type-2 manipulation.

[50] It might give me reasons to double check that the feeling is accurate, but not to revise it once I've established that it is.

[51] For an illuminating discussion of this point, see Moran 2002.

that I do not experience in cases of covert manipulation— and this might well be a further wrong I suffer.[52] But even so, this wrong is parasitic on the central wrong of having my practical agency coopted by the manipulator in the way described above.[53]

IV. MANIPULATION AND THE MEANS PRINCIPLE

This is thus the central wrong of manipulation: When we are manipulated, our practical agency is coopted by the manipulator and turned against itself. Because of the misalignment generated by the manipulator between our apparent reasons for action and our actual reasons, the more we try to exercise our practical agency well to adequately respond to our reasons for action, the less effective we are in doing so. I will now argue that this wrong is compounded by a further, related, one. Manipulation involves coopting our practical agency in this way *in order to induce us to pursue the ends of the manipulator*. Manipulation is thus also wrong because it involves using the manipulee as means.

Philosophers regularly invoke the means principle in debates concerning the permissibility of using someone's body or their suffering to achieve certain ends — for example, the permissibility of torturing someone to obtain information, punishing wrongdoers for the purpose of deterrence, or harming innocents to prevent a larger number of innocent people from being harmed. Manipulation involves a special

[52] A sadistic manipulator might not be happy to simply exploit my practical agency, they might also want me to be aware of that.

[53] In his account of the wrong of torture, David Sussman also appeals to the idea of turning the agency of the victim against itself, but this idea is articulated in a profoundly different way by him. (Unsurprisingly, since torture is profoundly different from the type of manipulation considered here.) For him, the pain of torture is experienced by us as 'the voice of our own body' demanding that we give in to the requests of the torturer. Because of that, '[our] suffering is experienced as not just something the torturer inflicts on [us], but as something [we] do to [ourselves,] as a kind of self-betrayal worked through [our] body and its feelings' (Sussman 2005, p. 21). It's in this sense that 'torture forces its victim into the position of colluding against himself through his own affects and emotions, so that he experiences himself as simultaneously powerless and yet actively complicit in his own violation (p. 4). The crucial idea here is that it's the *experience* of our own pain, where that pain embodies someone else's will, that makes us complicit in the wrong we suffer. In my account, by contrast, what renders us accomplices in the process of our own wronging is the fact that it's our very effort to exercise our practical agency well that condemns us to failing to do so, given the manipulator's intervention. This process is something that we can experience, but more often than not we don't. And even when we do, the experience does not play a role in explaining how our agency is turned against itself. That our agency is turned against itself is rather the object of our experience.

form of using someone as a means. Instead of using their body or their suffering, it involves using someone's capacity to exercise their practical agency to serve the ends of the manipulator. Although when we are manipulated, we do retain our capacity to respond to our reasons for action —that capacity is not extinguished, but exploited by manipulators—,[54] the way in which we do that is put at the service of the realization of the ends of the manipulator. After all, the point of shaping the circumstances in which we operate so that our apparent reasons for action fail to track our actual ones is that this enables the manipulator to exploit our capacity to respond to reasons in order to induce us to pursue ends that they have set for us.[55]

To be sure, manipulation (of the kind we're considering here) does not involve using physical force or coercion, unlike torture or punishment. But employing force or coercion is not essential to the idea of using someone as a means. The gist of the means principle is that since we have the capacity to set our own ends, we suffer a distinctive wrong when, without our consent, we are made to serve ends that we have not chosen for ourselves.[56] Why should it matter whether we are made to serve these ends by using physical force and threats rather than psychological mechanisms such as emotional pressure or deception? Indeed, the objection against being used as a means seems particularly strong in the case of manipulation. For not only are we being exploited to realize the ends of the manipulator but, as we have seen, we are active participants in this process, rather than merely passive ones (as in the case when it's only our body or our capacity to suffer that are being used). Insofar as manipulators co-opt our very capacity to decide what we have reason to do in the way described above, we are made to actively pursue, through the exercise of our own practical agency, their ends. In this way, we become accomplices in perpetrating the very wrong of which we are victims.

---

[54] As I see it, this is one of the main lessons to be drawn from Buss 2005 (see, especially, p. 215).

[55] An alternative formulation of the view that manipulating someone involves treating them as means appeals to the thought that when we are manipulated we cannot possibly consent to how we are treated by others, since consent require the option of dissenting, and this option is 'in principle ruled out' by manipulation (O'Neill 1989, p. 113; see also Korsgaard 1996, p. 139). But I'm persuaded by Buss' objections against this formulation of the view. (See also Pallikkathayil 2010.)

[56] Let me state the obvious and note that the notion of treating someone as a means has been interpreted in different ways. The interpretation I rely upon here is fairly standard in recent discussions of deontological constraints. (See, for example, Tadros 2015.) More technical discussions of how the notion is best understood in Kant's scholarship are above my paygrade, and certainly I don't intend to settle them here.

Now, there are cases of coercion that also present this feature. Unlike in the case of torture or punishment, when we comply with a threat, we are made to actively pursue someone else's ends. Still, manipulation seems especially wrong, all things equal, as compared to cases of coercion involving threats. This is because when we are threatened, new actual reasons are created for us, so as to ensure that our ends and the ends of the coercer coincide. For example, when I threaten to shoot your children if you don't open the safe, opening the safe becomes a genuine end of yours, since doing so is now instrumental to keeping your children unharmed. I have changed what you have, all-things-considered, reason to do, counting on the fact that, if you correctly grasp what you have reason to do, you will open the safe. In cases of manipulation, by contrast, the balance of reasons that determines what we ought to do remains unchanged. This is because no new actual reasons for action have been created, only apparent ones. In this sense, when we are manipulated, we pursue the ends of the manipulator despite the fact that they are not in fact also our own ends.[57] We do so either because we mistakenly believe that they are (as when Othello is misled to believe that Desdemona is unfaithful) or because our emotions and attitudes have been interfered with so that doing so feels like an appropriate response, when in fact it is not (as in the case where I make you feel nervous about the vaccine).

Importantly, this is also true in cases of paternalistic manipulation. When I manipulate you to ensure that you achieve a certain goal you want to pursue, I'm treating your happiness as my own end, but I am not treating you as a rational agent free to choose whether to pursue your own ends and how. I am still imposing my end on you. It's just that in this case my end is ensuring that you get to achieve the goal you are pursuing.

The way in which manipulation violates the means principle is thus especially egregious for two reasons. First, when we are manipulated, our own agency is exploited in the way discussed above. Thus, unlike in the case of torture or punishment, we ourselves play an active role (albeit often unwittingly) in bringing about the wrong that is being inflicted on us. Second, the way in which we are led to play such a role involves a failure on our part to appreciate what we have actual reason to do. The manipulator exploits our practical agency in order to induce a sort of alienation from our reasons for action that is absent in cases of coercion.

[57] For an excellent discussion of this point, see Yaffe 2003 (especially, pp. 337–41).

V. CONCLUSION

Manipulation is wrong for many reasons. Here, I have tried to vindicate the traditional idea that one of these reasons has to do with the fact that manipulation undermines our efforts to exercise our practical agency, understood as our capacity to respond to our reasons for action. I have done so by arguing that there is an important sense in which adequately exercising our practical agency is a matter of responding to our actual reasons for action, and not simply our apparent ones. If so, an obvious way in which manipulation wrongs us is by setting back our interest in adequately responding to these reasons.

This answer, however, is not enough to capture the distinctive wrong of manipulation, since the same interest can be undermined by other forms of interference. To understand the distinctive wrong of manipulation we need to consider the particular way in which this interest is set back when we are manipulated. Manipulation involves intentionally exploiting our capacity to exercise our practical agency in order to undermine our efforts to adequately respond to our reasons for action. When this is the case, there is a sense in which our practical agency is co-opted by the manipulator and turned against itself. Instead of directly controlling our agency, as when we are hypnotised, the manipulator shapes the circumstances in which we operate so that it is through our own decisions that we end up falling short in responding to the reasons for action that apply to us. Given the mismatch that has been intentionally (or at least recklessly) produced between our apparent reasons for action and the actual ones, the more we try to appropriately respond to what we have reason to do, the more we fail to do so. In this way, we become accomplices in the process of wrongful interference with our own practical agency.

Because this sort of interference is in the service of inducing the victim to pursue the ends of the manipulator, a further reason manipulation is wrong is that it violates the means principle. When we are manipulated, our capacity to exercise our practical agency is exploited, without our consent, to pursue ends that the manipulator has imposed on us. Indeed, to the extent that we are active participants in this process, manipulation involves an especially egregious violation of the means principle, as compared to cases in which it is our body or our suffering that is being used to pursue ends to which we have not consented.

ACKNOWLEDGMENTS

Thanks to Marcia Baron, Gabriel De Marco, Antony Duff, Helen Frowe, Todd Karhu, Gerald Lang and Dana Nelkin for helpful written comments. I'm also grateful to Sarah Buss, Ian Carter, Lars Christie, Tom Dougherty, Tom Douglas, Cécile Fabre, John

Martin Fischer, Michael Garnett, Jonathan Gingerich, Jan Kandiyali, Andrei Marmor, Michael Moore, Sophia Moreau, Rob Noggle, Serena Olsaretti, David Owens, Tom Pink, Arthur Ripstein, Joseph Raz, Geoff Sayre-McCord, David Shoemaker, Rob Simpson, Adam Slavny, Victor Tadros, Susan Wolf, Lorenzo Zucca, as well as to audiences at King's College London, UNC Chapell Hill, Cornell University, the University of Toronto, the University of Illinois Urbana-Champaign, the University of Inland Norway, the University of Warwick, a workshop organized by the Uheiro Centre at the University of Zurich and the 2022 Pavia Graduate Conference in Political Philosophy. Finally, I'm especially grateful to two anonymous reviewers for generous comments that greatly improved the paper, and to Bob Goodin for valuable editorial suggestions. The research for this essay was financially supported by a Research Fellowship of the Leverhulme Trust (RF-2020-351\8).

COMPETING INTERESTS

The author declares that he has no competing interests.

REFERENCES

Arpaly, Nomy. 2002. Moral worth. *Journal of Philosophy*, 99 (5): 223–45. https://doi.org/10.2307/3655647

Baron, Marcia. 2003. Manipulativeness. *Proceedings and Addresses of the American Philosophical Association*, 77 (2): 37–54. https://doi.org/10.2307/3219740

Baron, Marcia. 2014. The *mens rea* and moral status of manipulation. Pp. 98–120 in *Manipulation*, ed. Christian Coons and Michael Weber. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199338207.003.0005

Broome, John. 2021. *Normativity, Rationality, and Reasoning*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198824848.001.0001

Buss, Sarah. 2005. Valuing autonomy and respecting persons: manipulation, seduction, and the basis of moral constraints. *Ethics*, 115 (2): 195–235. https://doi.org/10.1086/426304

Cave, Eric M. 2007. What's wrong with motive manipulation? *Ethical Theory and Moral Practice*, 10 (2): 129–44. https://doi.org/10.1007/s10677-006-9052-4

Chang, Ruth. 2013. Grounding practical normativity: going hybrid. *Philosophical Studies*, 164 (1): 163–87. https://doi.org/10.1007/s11098-013-0092-z

Cohen, Shlomo. 2018. Manipulation and deception. *Australasian Journal of Philosophy*, 96 (3): 483–97. https://doi.org/10.1080/00048402.2017.1386692

Coons, Christian and Weber, Michael. 2014. Introduction. Pp. 2–16 in *Manipulation*, ed. Christian Coons and Michael Weber. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199338207.003.0001

Doris, John M. 2002. *Lack of Character*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139878364

Dworkin, Ronald. 2011. *Justice for Hedgehogs*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/9780674059337

Fischer, John Martin. 2004. Responsibility and manipulation. *Journal of Ethics*, 8 (2): 145–77. https://doi.org/10.1023/B:JOET.0000018773.97209.84

Gardner, John. 2007. The Gist of Excuses. Pp. 121–39 in Gardner, *Offences and Defences*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199239351.003.0006

Gibert, Sophie. 2023. The wrong of wrongful manipulation. *Philosophy & Public Affairs*, 51 (4): 333–72. https://doi.org/10.1111/papa.12247

Goodin, Robert. 1980. *Manipulatory Politics*. New Haven, CT: Yale University Press.

Gorin, Moti. 2014a. Do manipulators always threaten rationality? *American Philosophical Quarterly*, 51 (1): 51–61

Gorin, Moti. 2014b. Towards a theory of interpersonal manipulation. Pp. 73–97 in *Manipulation*, ed. Christian Coons and Michael Weber. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199338207.003.0004

Gorin, Moti. 2018. Paternalistic manipulation. Pp. 236–47 in *The Routledge Handbook of Paternalism*, ed. Kalle Grill and Jason Hanna. New York: Routledge. https://doi.org/10.4324/9781315657080-20

Haji, Ishtiyaque, and Stefaan E. Cuypers. 2008. *Moral Responsibility, Authenticity, and Education*. London: Routledge. https://doi.org/10.4324/9780203895146

Hanna, Jason. 2015. Libertarian paternalism, manipulation, and the shaping of preferences. *Social Theory and Practice*, 41 (4): 618–43. https://doi.org/10.5840/soctheorpract201541434

Hardimon, Michael O. 1994. *Hegel's Social Philosophy: The Project of Reconciliation*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511624773

Harman, Elizabeth. 2011. Does moral ignorance exculpate? *Ratio*, 24 (4): 443–68. https://doi.org/10.1111/j.1467-9329.2011.00511.x

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Korsgaard, Christine M. 1996. *Creating the Kingdom of Ends*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139174503

Levy, Neil. 2019. Nudge nudge wink wink: nudging is giving reasons. *Ergo*, 6 (10): 281–302. https://doi.org/10.3998/ergo.12405314.0006.010

Lewis, Paul and Rob Evans. 2010. Secrets and lies: untangling the UK 'Spy Cops' scandal. *The Guardian*, https://www.theguardian.com/uk-news/2020/oct/28/secrets-and-lies-untangling-the-uk-spy-cops-scandal

Markovits, Julia. 2010. Acting for the right reasons. *Philosophical Review*, 119 (2): 201–42. https://doi.org/10.1215/00318108-2009-037

Mele, Alfred R. 2019. *Manipulated Agents*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780190927967.001.0001

Mills, Claudia. 1995. Politics and manipulation. *Social Theory and Practice*, 21 (1): 97–112. https://doi.org/10.5840/soctheorpract199521120

Moran, Richard. 2002. Frankfurt on identification. Pp. 189–217 in *The Contours of Agency*, ed. Sarah Buss and Lee Overton. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/2143.003.0010

Noggle, Robert. 1996. Manipulative actions. *American Philosophical Quarterly,* 33 (1): 43–55.

Noggle, Robert. 2020. Pressure, trickery, and a unified account of manipulation. *American Philosophical Quarterly,* 57 (3): 241–52. https://doi.org/10.2307/48574436

Noggle, Robert. 2022. The ethics of manipulation. *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/sum2022/entries/ethics-manipulation.

O'Neill, Onora. 1989. *Constructions of Reason*. New York: Cambridge University Press.

Pallikkathayil, Japa. 2010. Rethinking the formula of humanity. *Ethics,* 121 (1):116–147. https://doi.org/10.1086/656041

Parfit, Derek. 2011. *On What Matters: Volume 1*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199572809.001.0001

Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685516.001.0001

Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.

Raz, Joseph. 2000. *Engaging Reason*. Oxford: Oxford University Press.

Raz, Joseph. 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199693818.001.0001

Renzo, Massimo. 2021. Manipulation and liability to defensive harm. *Philosophical Studies,* 178 (11): 3483–3501. https://doi.org/10.1007/s11098-021-01610-7

Renzo, Massimo. 2023. Responsibility and manipulation. Pp. 250–260 in *Routledge Handbook of Responsibility*, ed. Maximilian Kiener. London: Routledge. https://doi.org/10.4324/9781003282242-28

Renzo, Massimo. 2025. Motivation and alienation. Pp. 172–192 in *Engaging Raz: Themes in Normative Philosophy*, ed. Andrei Marmor, Kimberley Brownlee, and David Enoch Oxford: Oxford University Press. https://doi.org/10.1093/9780198925378.003.0010

Renzo, Massimo. Forthcoming. Manipulation and practical agency. *Free & Equal: A Journal of Ethics and Public Affairs*.

Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press. https://doi.org/10.2307/j.ctv134vmrn

Susser, Daniel; Beate Roessler; and Helen Nissenbaum. 2019. Online manipulation: hidden influences in a digital world. *Georgetown Law Technology Review*, 4: 1–45. https://doi.org/10.2139/ssrn.3306006

Sussman, David. 2005. What's wrong with torture? *Philosophy & Public Affairs,* 33 (1): 1–33. https://doi.org/10.1111/j.1088-4963.2005.00023.x

Tadros, Victor. 2015. Wrongful intentions without closeness. *Philosophy & Public Affairs,* 43 (1): 52–74. https://doi.org/10.1111/papa.12043

Thaler, Richard H., and Cass R. Sunstein. 2021. *Nudge: The Final Edition*. New Haven, CT: Yale University Press.

Wall, Steven. 2021. Self-mastery and the quality of a life. Pp. 102–19 in *Positive Freedom*, ed. John Christman. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108768276.007

Wood, Allen W. 2014. Coercion, manipulation, exploitation. Pp. 17–50 in *Manipulation*, ed. Christian Coons and Michael Weber. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199338207.003.0002

Yaffe, Gideon. 2003. Indoctrination, coercion and freedom of will. *Philosophy & Phenomenological Research,* 67 (2): 335–56. https://doi.org/10.1111/j.1933-1592.2003.tb00293.x