Manipulation

Patrick Todd

At the most general level, manipulation is one of many ways of influencing behavior, along with (but to be distinguished from) other ways, such as coercion and rational persuasion. Like these other ways of influencing behavior, manipulation is of crucial importance in various ethical contexts. First, there are important questions concerning the moral status of manipulation itself; manipulation seems to be morally problematic in ways in which (say) rational persuasion does not. Why is this so? Furthermore, the notion of manipulation has played an increasingly central role in debates about free will and moral responsibility. Despite its significance in these (and other) contexts, however, the notion of manipulation itself remains deeply vexed. I would say *notoriously* vexed, but in fact direct philosophical treatments of the notion of manipulation are few and far between, and those that do exist are notable for the sometimes widely divergent conclusions they reach concerning what it is. I begin by addressing (though certainly not resolving) the *conceptual* issue of how to distinguish manipulation from other ways of influencing behavior. Along the way, I also briefly address the (intimately related) question of the *moral status* of manipulation: what, if anything, makes it morally problematic? Then I discuss the controversial ways in which the notion of manipulation has been employed in contemporary debates about free will and moral responsibility.

What is Manipulation? What is its Moral Status?

We can start characterizing manipulation by contrasting it with coercion (*see* COERCION), on the one hand, and rational persuasion on the other. Suppose Cody issues the familiar threat to Valerie: your money or your life. Here we have a case of (attempted) *coercion*. Consider instead a case in which Rick, a charity worker, tries to elicit a donation from Valerie by pointing out (what he takes to be) the moral reasons she has to aid those in need. Here we have a case of (attempted) *rational persuasion*. Both Cody and Rick are trying to influence Valerie's behavior – both are trying to get her to hand over some money – but neither is attempting to manipulate her. Why not? One natural thought here would be the following. Both such ways of influencing Valerie are in some sense "overt"; in both cases, Valerie is fully aware of the ways in which Cody and Rick are trying to influence her. But manipulation is instead "covert" rather than "overt." That is, manipulation essentially involves *deception*; those who allege that they were *manipulated* claim that there were hidden factors at least partially responsible for their actions. As a first approximation, then, one might think to distinguish manipulation from coercion and rational persuasion as follows: to attempt to manipulate someone is to attempt to influence his or her behavior by means of deception.

Yet we'll want to be careful how we appeal to deception here. In particular, not all cases of influencing behavior by means of deception would seem to be cases of manipulation. Consider a case (slightly modified) from Rudinow, in which "a man stands at a fork in the road and knowingly directs to Minsk people who ask the way to Pinsk," perhaps because he gets a dollar for every car that ends up in Minsk (Rudinow 1978: 343). Intuitively, this does not rise to the level of manipulation. Why not? According to Rudinow, because the man is simply trying to influence the travelers' behavior in ways that take advantage of – rather than modify – their previously established goals, viz., getting to Pinsk. However, says Rudinow, if one tries to motivate someone's behavior by trying to change her *goals* (or by getting her to adopt new ones), and one does this by means of deception, then one is engaged in manipulation. But such a suggestion also faces problems. Suppose the man instead tells those who are on the way to Pinsk that a horrible accident has closed the road to Pinsk and so they had better go to Minsk instead. Here the man seemingly *has* changed their goals by means of deception, yet on finding out the truth, such travelers are likely to report feeling duped, but probably not manipulated. I suspect we are reluctant (if we are) to call such cases instances of manipulation at least in part because the given influence on behavior is not predicated on any special knowledge of the target's situation or psychology, but instead on the quite general fact that we are inclined to trust the testimony of others when there does not appear to be reason not to. In short, these tactics are not "sophisticated" enough to count as manipulation, and one task for an account of manipulation would be to explain the sort of sophistication at issue.

Consider instead another (or additional) way one might try to distinguish manipulation from coercion and rational persuasion. Instead of focusing on the ways in which manipulation is (allegedly) "covert," we might instead focus on the *capacities* of the targets of relevant kinds of influence – in particular, their capacities for rational reflection and deliberation. In cases of attempted coercion and rational persuasion, one does not try to *bypass* these capacities, but instead one tries to influence the target in ways that *go through* them (*see* MORAL AGENCY). (In cases of coercion, of course, one "goes through" such capacities only in the sense that one tries to *overwhelm* them; one tries to present one's target with considerations she will regard as irresistibly compelling compliance with one's demands.) One promising conception of manipulation is that it instead seeks to influence behavior in ways that (in some sense) *bypass* the target's capacities for rational reflection altogether. (For discussion of a similar notion of "bypassing," see Mele 1995.) One challenge for such an account would be to avoid construing manipulation so widely that various benign ways of trying to influence behavior turn out to be manipulative. Consider the wearing of perfume or cologne. If Jones wears cologne on a date, then he is likely trying to influence his date's behavior, but likely not by prompting her to rationally reflect on his choice to wear cologne. Rather, he is likely trying – perhaps pathetically – to get his date to simply *feel* attraction toward him, without ever having rationally reflected about why this would (or should) be so. Yet I think many will be reluctant to say that those (like Jones) who (say) try to appear attractive before going on dates are attempting to manipulate their dates.

Why is this so? Again, I suspect that we are reluctant (if we are) to call this sort of behavior attempted manipulation at least partially because such actions are again typically not predicated on any special knowledge of the target's situation or psychology. (Similar remarks could apply to various sorts of advertising and marketing; typically, marketers target certain groups *generally*, and are only able to exploit their knowledge concerning *general* facts about human psychology. This may explain why one might think that while various forms of advertising and marketing are morally problematic, they typically do not rise to the level of *manipulation*.) On the other hand, we may be reluctant to regard the wearing of cologne as attempted manipulation because, typically, engaging in such "tactics" is in accord with the commonly known and accepted "rules" governing such interpersonal interactions; when one agrees to go on a "date" one generally understands that one is entering into a social context in which such tactics may be employed. Thus, if one is swayed by such tactics, one cannot complain of having been *manipulated* by them, since one knew (or should have known) that they would probably be employed, and thus one should have taken precautions against them.

One operative assumption of the preceding discussion is that an adequate account of manipulation should not include ways of influencing behavior that we rightly regard to be morally unproblematic. Plausibly, so the thought would go, manipulation is at least *prima facie* wrong in the following sense: if one is attempting to manipulate someone else, then one's actions are permissible only if one has an *excuse* or a *justification* for doing so. One doesn't need an excuse or a justification for walking around the block, or, plausibly, for wearing cologne on a date, but it strongly seems that one *does* need an excuse for attempting to manipulate someone, the standard worry being that manipulation somehow violates or undermines *autonomy* (*see* AUTONOMY). Of course, some may deny that manipulation must even be *prima facie* wrong in this sense. Some, however, might argue that manipulation is not only always *prima facie* wrong, but always also *ultima facie* wrong – that is, that there can be no excuse or justification for it. Presumably, those theorists most likely to be attracted to this thesis will be those attracted to the (notoriously difficult) Kantian thesis that it is always and everywhere wrong to treat others as merely means to an end, together with the claim that manipulation always involves treating others in precisely this way (*see* KANT, IMMANUEL; CATEGORICAL IMPERATIVE). Notably, those taking this line would (minimally) have to explain why Jones' wearing of cologne is not manipulative, or they would have to explain how it is that in doing so he problematically treats his date "as a mere means."

Another issue worth discussing concerning the concept of manipulation – and one that will help us transition to the next section – concerns the relationship between manipulation *per se* and *acting manipulatively* or *being manipulative*. Arguably, those (few) who have been interested in the conceptual issues surrounding manipulation have only been after an account of what it is to *act manipulatively* (Rudinow 1978; Greenspan 1993; Noggle 1996). That is, in seemingly all of the cases of (putative) manipulation discussed in this literature, it seems appropriate to say

that the manipulator is *acting manipulatively* or *being manipulative*, where this involves the manifestation of the characteristic vice of *having a manipulative personality*. We might call any case of manipulation in which it is appropriate to say that the manipulator is *being manipulative* a case of *interpersonal manipulation*. However, there are plausibly cases of manipulation in which it would *not* be appropriate to say that the given manipulators are *being manipulative* or *acting manipulatively*, cases that (understandably) have not occupied the attention of those whose interests in manipulation stem from the ethics of interpersonal relationships. Of course, it can seem simply *analytic* – true simply in virtue of the meanings of terms – that any case of manipulation must also be a case of someone acting manipulatively or being manipulative, but this is a thesis we should (arguably) resist.

Consider, for instance, cases familiar to philosophers who work on the problems of free will (*see* FREE WILL) and moral responsibility (*see* RESPONSIBILITY), cases in which, say, some very powerful neuroscientists have covertly implanted a chip in Jones' brain that allows them to influence his behavior via remote control. Suppose the chip works by directly inducing various desires in Jones, or suppose it works by triggering the right release of chemicals in Jones' brain that will be sufficient for Jones' coming to have a given belief or desire (or …), and suppose the neuroscientists are doing this as part of some experiment in behavioral control. Now, it seems clear that these neuroscientists are manipulating Jones by means of the chip. However, it doesn't seem clearly right to say that they are being manipulative or acting manipulatively, anyway not in anything like the pejorative ("That manipulative jerk!") sense in which these terms are typically used. In particular, while they very plausibly have *some* vices, we needn't suppose at all that these neuroscientists exhibit manipulative personalities when engaged in *interpersonal* relationships with, say, friends or loved ones. The neuroscientists are here manipulating Jones in the sense in which an engineer might manipulate the controls on a machine, and it would clearly be odd to say of such an engineer that she is, on this account, "so manipulative." This is because the engineer is not engaged in any sort of interpersonal relationship with what she manipulates. And neither are the neuroscientists engaged in any such relationship with Jones (however different he may be from a machine). The upshot here is this. When we're theorizing about manipulation, we ought to take care to distinguish manipulation *per se* and cases in which agents are acting manipulatively or being manipulative.

A final word is in order here concerning the concept of manipulation. After considering the issues of this section, one might naturally think that there is no such single thing that deserves to be called "manipulation," and that there simply are no informative necessary and sufficient conditions for something's being a case of manipulation. Rather, the concept of manipulation is best thought of as a "family resemblance" concept. On this view, various things might deserve to be called "manipulation," and while there will be resemblances between all such cases, there will not be some one core unifying (and informative) property they all have in common. Given the widely divergent contexts in which the notion of manipulation seems appropriately used, this view can seem particularly plausible.

Manipulation, Free Will, and Moral Responsibility

Perhaps the easiest way to begin to see how the notion of manipulation has figured into debates about free will and moral responsibility is to note the ways in which it has caused problems for one account of the conditions under which agents are free and responsible. (For simplicity, I will not distinguish here between the conditions for free will and moral responsibility, if these conditions differ.) Consider the sort of "mesh" view defended by Harry Frankfurt (1971; Fischer 1994: 208–9). On this view, it is sufficient for acting freely and responsibly that (roughly) one acts from desires that exhibit the right "mesh" with one's higher-order desires. For instance, suppose that not only does Jones want to rob the store (a first-order desire), but also he wants to want to rob the store (a second-order desire) – he *endorses* being the sort of person who wants to rob stores, so that robbing stores expresses (in some sense) "where he stands." Then, if his desire to rob the store culminates in his doing so, he did so freely and responsibly. The crucial idea here is that it is enough for freedom and responsibility that we act from the "wills we want to have."

But suppose Jones is the same Jones considered above, who is subject to the control of the neuroscientists. Suppose they directly induce in Jones the desire to rob the store. But they also want to bring it about that Jones *freely* and *responsibly* robs the store, so, heeding Frankfurt, they similarly bring it about that Jones wants to want to rob the store. In such a scenario, Frankfurt would be committed to saying that Jones robs the store freely and responsibly (and thus perhaps deserves blame or even punishment), but this can seem like the wrong result (*see* BLAME; DESERT; PUNISHMENT). The problem is that I could be acting from a will I want to have, but it could nevertheless be that I was manipulated into wanting to have that will. In other words, we shouldn't pay attention *only* to features of the agent at the time of action, e.g., whether her action was in accord with her second-order desires, and so on. Manipulation scenarios (arguably) show that we should also pay attention to how the agent got to be that way in the first place; perhaps she came to be that way via responsibility-undermining manipulation. In short, cases involving manipulation seem to be counterexamples to Frankfurt's account of free will and moral responsibility.

So cases involving manipulation have posed problems for some accounts of free will and moral responsibility. Some philosophers, however, have gone on to argue that similar cases in fact undermine a whole *class* of such theories – namely, *compatibilist* theories on which free will and moral responsibility are compatible with *determinism* (Kane 1996; Pereboom 2001). The doctrine of determinism can be hard to understand, and how it should be defined is controversial. Roughly, however, determinism is the thesis that past states of the world, together with the laws of nature, entail a unique future. More concretely, if determinism is true, then all of our behavior is (in some sense) "fixed" by such factors as our genes, upbringing, and environment. In principle, then, if someone knew enough about the laws of nature, and knew enough about our pasts, then she could predict with perfect accuracy everything that we ever do. And compatibilism is the thesis that free will and moral responsibility are compatible

with determinism, so understood. Now, at least one way of putting the challenge for compatibilism is as follows. If determinism is true, then there are law-like connections between certain prior states of the world and certain human actions. Thus, it seems, we could imagine certain powerful manipulators who know about these sorts of law-like connections, and are thus able to guarantee in advance everything we do. But it can seem that anyone that has been manipulated or controlled in this sort of way could be neither free nor responsible. Compatibilists, however, seem committed to the claim that some such agents could indeed be free or responsible, depending on the details of how these agents were deterministically "set up" to perform the relevant actions. And some take this to be a cost for compatibilism.

Of course, the compatibilist will again insist that it all depends on the details. If the manipulation unfolds in some ways, then the compatibilist could plausibly insist that the relevant agent does not in fact meet plausible compatibilist conditions for responsibility. Thus, according to the compatibilist, there would be some freedom and/or responsibility relevant difference between an agent manipulated in this sort of way, and otherwise "normal" agents in a deterministic world. This would be to take a "soft-line" reply to the given case (McKenna 2008). To take a "hard-line" reply to the case, on the other hand, would be to admit that the given agent is indeed free and responsible, despite having been "set up" in the relevant way, but to contend that this is not an ultimately unacceptable result. Notably, however, the compatibilist is seemingly committed to taking a "hard-line" reply to at least *some* suitably formulated manipulation scenario, unless she contends that the fact that one's actions deterministically "trace back" to an intentional cause *in itself* rules out freedom and responsibility. One challenge for such a view – and an important issue more generally about the notion of manipulation – would be to explain what relevant difference there might be between the causes of one's actions being put in place by intentional agents, and (otherwise similar) causes having been there as a matter of brute natural chance. Of course, only *agents* can (strictly speaking) manipulate us, but couldn't nature itself be the functional equivalent of a manipulator?

In sum, manipulation scenarios have primarily been employed in order to challenge various compatibilist accounts of free will and responsibility. But work on these topics is certainly far from complete.

***See also*:** AUTONOMY; BLAME; CATEGORICAL IMPERATIVE; COERCION; DESERT; FREE WILL; KANT, IMMANUEL; MORAL AGENCY; PUNISHMENT; RESPONSIBILITY

REFERENCES

Fischer, John Martin 1994. *The Metaphysics of Free Will*. Oxford: Blackwell.
Frankfurt, Harry 1971. "Freedom of the Will and the Concept of a Person," *Journal of Philosophy*, vol. 66, pp. 829–39.
Greenspan, Patricia 1993. "The Problem With Manipulation," *American Philosophical Quarterly*, vol. 40, pp. 155–64.

Kane, Robert 1996. *The Significance of Free Will*. New York: Oxford University Press.

McKenna, Michael 2008. "A Hard-Line Reply to Pereboom's Four-Case Manipulation Argument," *Philosophy and Phenomenological Research*, vol. 77, pp. 142–59.

Mele, Alfred 1995. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.

Noggle, Robert 1996. "Manipulative Actions: A Conceptual and Moral Analysis," *American Philosophical Quarterly*, vol. 33, pp. 43–55.

Pereboom, Derk 2001. *Living Without Free Will*. New York: Cambridge University Press.

Rudinow, Joel 1978. "Manipulation," *Ethics*, vol. 88, pp. 338–47.

FURTHER READINGS

Cave, Eric M. 2007. "What's Wrong With Motive Manipulation?" *Ethical Theory and Moral Practice*, vol. 10, pp. 129–44.

Haji, Ishtiyaque, and Stefaan E. Cuypers 2008. *Moral Responsibility, Authenticity, and Education*. New York: Routledge.

McKenna, Michael 2004. "Responsibility and Globally Manipulated Agents," *Philosophical Topics*, vol. 32, pp. 169–92.

Mele, Alfred 2005. *Free Will and Luck*. Oxford: Oxford University Press.

Mele, Alfred 2008. "Manipulation, Compatibilism, and Moral Responsibility," *Journal of Ethics*, vol. 12, pp. 263–86.

Todd, Patrick 2011. "A New Approach to Manipulation Arguments," *Philosophical Studies*, vol. 152, pp. 127–33.

Watson, Gary 1999. "Soft Libertarianism and Hard Compatibilism," *Journal of Ethics*, vol. 3, pp. 351–65.