Philosophy and Phenomenological Research
Vol. LXVII, No. 2, September 2003

Indoctrination, Coercion and Freedom of Will

GIDEON YAFFE
University of Southern California

Manipulation by another person often undermines freedom. To explain this, a distinction is drawn between two forms of manipulation: indoctrination is defined as causing another person to respond to reasons in a pattern that serves the manipulator's ends; coercion as supplying another person with reasons that, given the pattern in which he responds to reasons, lead him to act in ways that serve the manipulator's ends. It is argued that both forms of manipulation undermine freedom because manipulators track the compliance of their victims, while neutral causal mechanisms do not. Manipulators see to it that their victims comply even in the face of forces that threaten to derail them from the manipulator's desired course. It is suggested that this has an impact on freedom because part of what we desire in wanting to be free is the availability of forms of life very different from those we actually enjoy.

The fact that coercion and other forms of manipulation can undermine our freedom is difficult for any theory of the nature of freedom to account.[1] The problem is that it is very difficult to specify any sense in which manipulation by another closes or limits options that is not also a sense in which forces that are not usually thought to be freedom-undermining close or limit options. We need to explain why it is that freedom is sometimes undermined by manipulation by another when it would not be undermined by other causal forces that produce precisely the same actual physical and psychological effects on its victim.

Consider an example of this: Imagine two agents both of whom are trying to decide whether to give $1000 to Amnesty International or to David

[1] Freedom, in the sense under consideration here, is the sort that is necessary for moral responsibility. However, "morally responsible" is a rather ambiguous term, picking out various concepts that may or may not be co-extensive; sometimes the term refers to those who deserve punishment or reward, sometimes to those who deserve blame or praise, sometimes to those who are invested with capacities for the appreciation of moral value, etc. These ambiguities bleed into our intuitions about which agents are and are not morally responsible for what they've done. This implies that efforts to assess a theory of freedom by consulting intuitions about moral responsibility rest on dangerously shifting sands. If we must rely on intuitions about the application of concepts in order to assess a theory of freedom, why not just rely on intuitions about how to apply the concept of freedom itself?

Koresh. The first, determined to make her money do as much good as possible, weighs the pros and cons of each choice and decides to give her money to Koresh since her tea leaves and sheep guts, which have never let her down before and which she has taught herself to use as well as they can possibly be used, tell her that he is the messiah. The second is singled out by Koresh's minions and lured to his compound for a weekend "seminar" at the conclusion of which she believes Koresh to be the messiah. She then deliberates carefully about what to do and decides to give her money to Koresh, since, after all, she believes him to be the messiah. Both agents, we can assume, are making the wrong choice—their money would do much more good in the hands of Amnesty International—but the first, and not the second, seems to be choosing freely. Why? Both agents are subject to forces that make them believe that Koresh is the messiah, and both of them take that to be sufficient reason to give him their money. In the first case, included in the cause of the agent's belief is the particular configurations of sheep guts and tea leaves that happen to have occurred at a particular moment; in the second, the cause of the belief includes the machinations of various people who are aiming at producing the belief. But why does that difference matter?

We can't find the difference in the two cases merely by pointing to the fact that the second involves manipulation and the first does not, because that is simply to assert what we want to explain: namely, that manipulation seems, in certain circumstances, to interfere with freedom of will. It is true that in the second case, and not in the first, the will of another person is involved in producing the agent's choice. The question is what is significant about that fact? Why should the will of another have a negative impact on freedom? We cannot answer this question through appeal solely to the actual effects that each of the respective causal processes has on the agents in the examples, because the effects are precisely the same: both have the same belief and are disposed to respond to that belief in the same way. So whatever it is about the will of a manipulator that undermines freedom, it cannot always be duplicated by a causal mechanism that has just the same actual effects on the agent as the manipulator has. And this is a puzzle. If manipulation doesn't undermine freedom because of its actual effects, why does it undermine freedom?

It is possible to respond that, in fact, manipulation doesn't undermine freedom. Nothing to be said here rules out this response entirely. However, somebody who responds this way takes on the burden of explaining the mistake that many rational people well-versed in the ins-and-out of the free will debate must be making. That is, at the very least, the *intuition* that manipulation can undermine freedom needs to be explained, even if the explanation does not paint the intuition as veridical.[2] But even an explanation for the

[2] Not just any explanation for the intuition will do. In particular, attempts to explain the intuition while denying that it is veridical must nonetheless respect its contours. It is a fail-

intuition fails to arise immediately from either compatibilism or incompatibilism.

Traditionally, compatibilists have seen freedom as exhibited in an agent's action just in case her action is causally dependent on certain mental states, or complexes of mental states, that are thought to be, for various reasons, crucial for moral responsibility or other forms of assessment to which we subject only free agents. In its simplest form, on this kind of view, we are free just in case what we do depends on what we choose.[3] Since our choices and subsequent conduct can be deterministically caused, and we can still satisfy this condition, under this view, we can be free even under determinism. But as a general rule we seem unfree when our choices are the product of manipulation, even if the subsequent actions do depend on our choices. And so it seems that views of this sort can only explain the intuition that manipulation undermines freedom by saying that those who think so are simply mistaken about the nature of freedom. An explanation of this sort is not satisfactory, however, for the intuition doesn't go away after one has been instructed by compatibilists of this sort as to the nature of freedom.

Of course, this form of compatibilism is simplistic. However, the problem cannot be solved just by appealing to complexes of mental states, rather than the single mental state of choice. Say, for instance, we were to say that an agent is free just when her conduct depends on what she chooses and what she chooses depends on which of her desires, or sets of desires, she identifies with.[4] We are still faced with the possibility that a sophisticated manipulator

---

ure to do so that mars the following explanation of the intuition: We might say that the thought that manipulation undermines freedom is really just a misrepresentation, or misexpression, of a different, yet true, thought: namely, that those who are manipulated into choosing to act in some way are wronged. Such victims, we might say, are used, exploited, or treated as means rather than ends. On this line of thought, the victim's freedom is not actually undermined by manipulation, rather, she is treated as though she lacked the distinctively human features of agency, and, therefore, treated as though she were unfree. (In a later footnote, I discuss, and reject, a view under which such treatment does, by its very nature, undermine freedom.)

The trouble with this explanation is that the intuition doesn't alter under new information as it would were it really a misexpression of the thought that the victim is being mistreated. If I find out how X is aiming to manipulate Y, I now know that X is treating Y as a means; but if I now find out that Y doesn't respond to X's attempts at manipulation, I no longer think that Y is made unfree by X's efforts, even though I still think that Y is being mistreated. So, the two thoughts behave quite differently.

3 This view was held by Thomas Hobbes. (cf. Hobbes, "Of Liberty and Necessity" in *Hobbes and Bramhall on Liberty and Necessity*, Vere Chappell, ed. Cambridge University Press, Cambridge, 1999, pp. 15-42.) It was also held in a modified form by A. J. Ayer. (Ayer, "Freedom and Necessity", in *Free Will*, Gary Watson, ed. Oxford University Press, Oxford, 1982, pp. 15-23.)

4 Harry Frankfurt's conception of free agency is closest to the conception used here as an example. See, for instance, "Freedom of the Will and the Concept of a Person", in *The Importance of What We Care About*, Cambridge University Press, Cambridge, 1988, pp. 11-25. Frankfurt develops the problematic concept of identification in "Identification and

would not just induce the agent to choose in a particular way, but would also induce the agent to identify with her choice. Still, even once we fully understand such a compatibilist's conception of freedom, it seems that such an agent is not acting freely in all the senses in which we hope ourselves to be capable. In fact, such manipulation will seem to some to be even worse than more direct forms. If you are stuck with a companion who regularly gets you to do what she wants you to do, would you rather she were someone who gets her way by leading you to ignore those aims and wants with which you identify and do as she wants, or someone who perniciously works on you to alter what you identify with so that it conforms with what she wants you to care about? The second kind of manipulation has lasting effects that the first does not. In the first case, you can look back on your conduct later and see that you had no reason of your own to do as you did, while in the second, the fact that you were being used is, in effect, concealed from you: you will see yourself later—and you won't be mistaken—as having had reasons of your own for doing as you did. I count myself as among those who are frightened by the second sort of companion more than the first: if I have to choose, I would rather be a pawn than a toady.

To be a pawn is to act as someone wants you to act while lacking some of the attitudes towards your action (or relations between attitudes) that are generally present when your actions are free. Such a person feels the force of the other's manipulation as doing violence to her agency; the manipulator causes her to lose a battle to the motives on which the manipulator wants her to act. To be someone's toady, on the other hand, is to be seduced; such a person doesn't lose a battle to the motives on which the manipulator wants her to act, rather, the manipulator induces her to side with them. To be a toady is to be free to yourself, and yet not to be free. Such an agent possesses all of the subjectively accessible mental states and attitudes that are involved in free action and those attitudes all play just the roles in the production of her action that they play when she acts freely, but, nonetheless, she is not (or, at least, seems not to be) free: she feels that she acts freely only because her manipulator makes her feel that way. What we need, I believe, is an explanation for the fact that it seems—despite the opposing view of the two kinds of compatibilism mentioned—perfectly possible to be a toady. And it is such an explanation that I aim to give here.

Such an explanation is no easier for incompatibilists to provide than it is for compatibilists. One might think that having explained why it is that freedom of will is undermined by determinism, we would thereby have explained

---

Externality", in *The Identities of Persons*, Amélie Rorty, ed. University of California Press, Berkeley, 1976; "Identification and Wholeheartedness", in *Perspectives on Moral Responsibility*, John Fischer and Mark Ravizza, eds. Cornell University Press, Ithaca, 1993, pp. 170-187; and in "The Faintest Passion" in *Necessity, Volition and Love*, Cambridge University Press, Cambridge, 1999, pp. 95-107.

why freedom is undermined by manipulation even while remaining neutral on the question of whether or not anyone is actually free. Manipulation, an incompatibilist might say, undermines freedom because it is a form of causal determination, something which the incompatibilist takes, on independent grounds, to undermine freedom.[5] But since sticks and stones aren't free even if determinism is false, it follows that it cannot be plausibly claimed that the mere falsity of determinism is sufficient for freedom. Intuitively it appears that manipulation can undermine freedom even if determinism is false, and so manipulation seems to undermine freedom by falsifying some other condition required for freedom beyond the falsity of determinism. Say, for instance, that a manipulator can only make me choose as she wishes with ninety-nine percent certainty. Occasionally, if there is a random blip at just the right moment, her manipulative efforts fail. Still, if I do choose as she wishes it seems that I don't choose freely. The point is really a familiar one from the free will literature: when someone chooses so as to comply with the aims of a manipulator, we can cite her manipulation as an excuse for her conduct without showing (how would we show it?) that determinism is true. The problem with manipulation is not that it is a form of causal determination (although it might be).

Of course, an incompatibilist can be a libertarian and thereby supply necessary (and sufficient) conditions for freedom that can only be satisfied in the absence of both determinism and manipulation.[6] While such an account might be extensionally adequate—the libertarian might give conditions for choosing freely which can only be satisfied by those who are not victims of manipulation—this is no guarantee of providing an explanation for the freedom-undermining force of manipulation. Say that every time an agent makes a free choice, that choice is not determined by prior causes, but the agent adds something of her own. And say that something, whatever it is, which the agent adds is absent in all cases of manipulation. From here libertarian accounts might differ from one another: under some accounts all cases of free-

---

5 Another kind of incompatibilist reverses the order of explanation. She takes it for granted that manipulation undermines freedom and then argues that any explanation for this fact will also catch determinism in the net of freedom-undermining influences. This argument assumes that there is no explanation for the freedom-undermining effect of manipulation that validly distinguishes manipulation from neutral causal forces. Although this paper provides an explanation that distinguishes between the two, nothing said here establishes that the distinction identified is anything more than a source of conceptual confusion resulting in a mistaken intuition that there is a difference between manipulation and neutral causal forces. See the remarks towards the end of the section entitled "Indoctrination" for an elaboration of this point.

6 A recent, prominent libertarian position is Robert Kane, *The Significance of Free Will*, Oxford University Press, Oxford, 1996. Also of importance is the libertarian position offered in Randolph Clarke, "Toward a Credible Agent-Causal Account of Free Will", in *Agents, Causes and Events: Essays on Indeterminism and Free Will*, Timothy O'Connor, ed. Oxford University Press, New York, 1995, pp. 201-215.

dom-undermining manipulation are also cases of causal determination; for others, manipulation can undermine freedom even if it does not impose causal determination. Accounts of the former sort provide the same explanation for the fact that manipulation undermines freedom as they do for the fact that determinism does. Thus, they find themselves in the same situation as the incompatibilist neutral on the question of whether or not we are free: they lack a manipulation-specific explanation for its freedom-undermining effects and are thus overlooking the fact that our appeals to manipulation as evidence for the absence of freedom of will are not veiled appeals to the presence of causal determinism. On the other hand, if the libertarian explains the freedom-undermining effect of manipulation independently of the truth or falsity of determinism, then her explanation can be accepted by the compatibilist, and as we have seen, no appropriate explanation arises immediately from that quarter either. In short, then, we don't have a theory, and we need one.

Two Kinds of Manipulation

The effective manipulation of another—that is, manipulation which actually leads to the other choosing to do the manipulator's bidding—can be divided into at least two types. In both, the victim of manipulation sees herself as having most reason to do that action that the manipulator wants her to do (although she may or may not know that the manipulator wants her to do it). But sometimes this happens because the manipulator causes the victim to respond to reasons in a way which is advantageous to the manipulator, and sometimes this happens because the manipulator causes the victim to have reason to do only those things that it is advantageous to the manipulator for her to do. That is, sometimes we manipulate the way an agent responds to reasons, and sometimes we manipulate what reasons she has. The first kind of manipulation could be called indoctrination, the second coercion. For instance, a cult leader might manage to convince the new followers to give him, say, fifty percent of their annual incomes so as to promote his welfare, perhaps because he convinces them that he is worthy of their money while they are not. They may or may not have reasons to do his bidding—perhaps he really is more worthy than they—but his manipulation is effective regardless of the facts about what they have reason to do. By manipulating them, he causes them to recognize and respond to those features of the world which give *him* reason to have them act in particular ways. That is, he indoctrinates them. On the other hand, a robber, by holding a gun to the head of a cashier, effectively manipulates the cashier to choose to hand over cash. The robber doesn't alter the way in which the cashier responds to reasons. Rather, the robber relies on the cashier to respond to reasons as most of us do—self-interestedly—and effectively manipulates her by giving her overwhelming

self-interested reasons to choose to do as he wants her to. That is, he coerces her.

You can be a toady as the result of either indoctrination or coercion. While it is clear that the indoctrinated need not report any psychic conflict over their choices in order to be victims, the same is true of the coerced. Some cases of coercion, in contrast, involve irrational response to overwhelming fear—victims of this kind of coercion might feel, in retrospect, that they didn't act from the motives with which they identified. They might, that is, conceive of themselves as having acted from a complex of psychic states quite different from those ordinarily involved in free action. But this is not necessary. A cashier who coolly and calmly calculates the gains and losses of defiance and chooses to comply is no less a victim of coercion. Such a cashier is, for a time, the toady of the robber: the robber has made her do his bidding not by making her choose contrary to her own inclinations, but by making his interests coincident with hers. The cashier is as free to herself as any of us are when we make a decision in favor of self-interest, but she is not, nonetheless, free.[7]

In fact, the best explanation for the freedom-undermining force of both kinds of manipulation turns out to be the same, although the route to the explanation is slightly different in each case. I will consider each case in turn, but, first, a few remarks about "reasons" for action.

Reasons

The notion of someone's "reasons" for acting in a particular way is extremely slippery. In one sense, an agent has genuine reasons to act certain ways, given her beliefs, even if her beliefs are false. To take the earlier example, in causing someone to believe that Koresh is the messiah, Koresh gives her a genuine reason, relative to her beliefs, to give him her money. However, there is another sense in which such a person lacks any reason whatsoever to give Koresh her money: her belief that he is the messiah is false, and she only has genuine reason, in this second sense of the term, to give him her money if he really is the messiah. The first sense of the term "reasons" pertains only to the explanation of action, the second pertains, in some way or another, to the good. Let's say that someone who has a reason to act in the first sense has a "rationale" for acting as she does, and reserve the term "reason" for the second sense. In the second sense of "reasons", a person has reasons for performing a particular action just in case the performance of that action would actually be good in some way or another (which isn't to say it

---

7 Discussions of the impact of coercion on freedom often overlook the fact that one can be a toady as a result of coercion. Cf. Kane, *The Significance of Free Will*, pp. 134-135, and Michael Slote, "Understanding Free Will" in *Journal of Philosophy*, 1980, pp. 138-139. Both Kane and Slote seem to think that an agent who is made unfree by coercion suffers some kind of internal conflict.

would be good all things considered). Reasons, in this second sense, confer some degree of legitimacy on the actions they favor.[8]

When we are considering the freedom-undermining effects of indoctrination, we are interested only in reasons and not in rationales. Victims of indoctrination don't, generally, have any less of a rationale for their choices than they had for the choices which they made prior to the influence of the indoctrination. Indoctrination, in the rather technical sense in which I am using the term, affects how agents respond to reasons. In addition, it also tends to provide agents with reasons as well, since it provides them with new desires that they formerly lacked. The desire to give money to the messiah might have been implanted by the indoctrination together with the belief that the indoctrinator is the messiah. However, in so far as indoctrination provides agents with new reasons, it is having a coercive effect rather than the effect that is distinctive of indoctrination. The feature of indoctrination that interests us is this: victims of indoctrination evince a new pattern of taking facts to be reasons for acting in particular ways. (As a general rule, this pattern serves the ends of the manipulator.) It is safe to think of this pattern as in part a pattern of belief about which features of oneself and one's environment really are reasons for acting in particular ways, and in part a pattern of desiderative attitudes, or affective states, that move one to act in ways consistent with one's beliefs about which facts are reason-giving. These beliefs can be correct or incorrect depending on whether or not the features believed to be reason-giving really are reason-giving in some respect, and the accompanying desiderative attitudes can be consistent or inconsistent with the relevant beliefs.[9]

---

8 Under a Davidsonian model of human action, desires generally—but perhaps not always (because they could be self-destructive, impulsive, or fleeting)—give one both a rationale and a reason to act in ways that one believes will contribute to their satisfaction. For Davidson, desires, together with certain beliefs, always provide a rationale for conduct. (Donald Davidson, "Actions, Reasons and Causes" in *Essays on Actions and Events*, Clarendon Press, Oxford, 1980, pp. 3-19.) But desires provide a reason for acting certain ways, in the second sense of "reason", only when our beliefs about which actions will contribute to the satisfaction of our desires are true. Desire satisfaction is usually (although perhaps not always) a good, even if it is doesn't necessarily make that action that satisfies it good all things considered. If I have a desire to look at a corpse, or a pornographic image, then I have some reason to do so, even if I have stronger reasons not to. By doing it I gain desire-satisfaction, a good (perhaps), and I live a seedier life, an evil (surely).

9 The distinction between reasons and rationales is intended, at least, to be neutral with respect to the question of whether or not there are only "internal" reasons for action in the sense discussed by Bernard Williams (Bernard Williams, "Internal and External Reasons" in *Rational Action*, R. Harrison, ed. Cambridge University Press, Cambridge, 1979, pp. 17-28). That question concerns whether or not an agent only ever has a reason, in my sense of the term, to do those things that can be reached through some appropriate "deliberative route" from her set of desiderative states (her "S", to use Williams's term). Nothing about the notion of an internal reason precludes an agent from having an internal reason to do something that she has no desire to do, or something that she falsely believes to be of no use for the satisfaction of her desires. At least, nothing precludes either of

Armed with the distinction between reasons and rationales, we can return to the question of what it is about manipulation that undermines freedom. First, consider indoctrination.

Indoctrination

Let's call an agent who is disposed to evince a particular pattern of responsiveness to reasons as a result of manipulation "the Manipulated", and let's imagine another agent, call her "the Unlucky", who is just the same as the Manipulated, except in the place of the manipulator, who aims to have her evince a particular pattern of response to reasons so as to further his own nefarious aims, we imagine a causal force that is entirely indifferent to what pattern of reasons she evinces, but happens, nonetheless, to produce in her a choice-producing mechanism (the set of dispositions underlying the relevant beliefs and desiderative attitudes) with just the same features as that produced in the Manipulated agent. What is the difference between these two agents that accounts for the difference in their freedom?

The crucial fact about a manipulator who aims to produce in you a certain pattern of response to reasons is that he tracks the production in you of that pattern of response. It is true of the Manipulated, and not of the Unlucky, that were the Manipulated to stray in some way or another from coming to have dispositions to recognize and respond to reasons of the sort that the manipulator wants her to have, he would take steps to see to it that she was placed back on course. The Unlucky, on the other hand, would simply stray from the course and come to have a different pattern of response to reasons

---

these possibilities while the notion of a "deliberative route" is left unanalyzed. If an agent desires X and does not desire Y, she still might have an internal reason to Y because there is some deliberative route to Y from her desire to X. Similarly, there may be some deliberative route from her desire to X to Y even if she falsely believes that Y-ing would be of no use for satisfying her desire to X. Independently, then, of a substantive characterization of a "deliberative route", or the conditions under which it is correct to say that an agent has such a route available to her, the distinction between reasons and rationales does not prejudice the case against those who hold, like Williams, that all reasons are internal.

Notice, however, that in allowing for the conception of reasons described here and for the distinction between indoctrination and coercion we are implicitly rejecting egoistic subjectivism about value. That is, if what an agent has reason to do is solely a function of her desires or other desiderative attitudes, then it might be impossible to influence how an agent responds to reasons without thereby influencing what reasons she has; to influence how an agent responds to reasons, on such a view, is to change what reasons she has since the very mental elements that determine how she responds to reasons are also responsible for determining what reasons she has. I am assuming, then, that what an agent has reason to do is not merely a function of facts about the mental states and dispositions through which she responds to reasons. Since not every characterization of what it is to possess a deliberative route from some set of desiderative states to a particular action commits an internal reasons theorist to egoistic subjectivism about value, it is possible to remain neutral with respect to the debate over internal reasons while still firmly rejecting egoistic subjectivism.

from that of the Manipulated. Call the choice-producing mechanism shared by the Manipulated and the Unlucky, M. There are certain counterfactuals that are false of the Manipulated and true of the Unlucky. Some counterfactuals of the form "If circumstances C were to occur, then S would no longer have choice-producing mechanism M" are true when S is the Unlucky and false when S is the Manipulated. So, imagine a Kaspar Hauser-type person who is confined to a dark room from a very early age by someone determined to make him entirely docile and not prone to conduct aimed at overthrowing a particular monarch. The relevant Unlucky agent is someone who is similarly confined, and similarly conditioned, but not by a person who is aiming to condition him in this way but merely through misfortune of some sort. If some good-hearted person who the manipulator is perfectly capable of preventing were to discover the Manipulated's situation at an early stage, the manipulator would prevent the Manipulated's rescue. But were the good Samaritan to discover the Unlucky's situation, the Unlucky might very well grow up to overthrow the monarch; that is, the good Samaritan would derail production of the actual choice-producing mechanism of the Unlucky, but not of the Manipulated. There is a very real sense, then, in which manipulators limit our options in ways that neutral causal forces do not. And it is that sense in which manipulators take away from our freedom.

Let's put the point a slightly different way. In considering what an agent could or could not do—as we do when we assess her moral responsibility for what she did—it is natural to examine a set of possible worlds in which the actual mechanism through which the agent acted is held fixed. But, our choice-producing mechanisms are not essential to us. Persons are things that endure through changes even in the dispositions and abilities through which they recognize and respond to reasons. This implies that our fundamental dispositions for recognizing and responding to reasons can be afflictions, even if they are such as to preclude our recognition of them as afflictions. We can look at others (and sometimes at ourselves) and consider the question of whether they might have been more free had they not been influenced by those events and processes through which they came to respond to reasons as they do. When we look at an agent in this way, our minds range over a set of possible worlds in which the agent has different dispositions for recognizing and responding to reasons. But sometimes we consider these kinds of possibilities while holding fixed certain crucially formative aspects of the agent's actual history. Imagine someone, for instance, who grew up in poverty, worked her way into monetary success, and now runs a successful business in a coldly Machiavellian way, firing those, for instance, who are forced to miss work in order to deal with poverty-induced hardships of their own. We might wonder what she would have been like—kinder-hearted, perhaps—had she had a parent who tempered personal financial ambition with empathy for the

struggles of others. That is, we consider possible worlds in which new forces were present in the formation of the agent's dispositions for recognizing and responding to reasons. When we wonder about such things, we wonder about what kinds of lives were available to others (and ourselves) given the unchangeable facts about them and their histories: perhaps there is something about this businesswoman that would have led to her being what she is, given the poverty in which she grew up, even in the face of counteractive influences; or perhaps not.

Questions like this are not idle. They are questions not about which actions are available to agents at particular times, but, rather, about what kinds of lives, what kinds of pattern of response to reasons, were available to agents given their unchangeable features and the crucial features of their pasts. When we fall into the hands of indoctrinators, fewer lives are available to us than are available to us when we are simply the unlucky victims of neutral causal forces. (More about the term "available" below.) When the kind of life an agent leads—the kind of pattern of response to reasons that she evinces—is the product of manipulation, then reveries about what the agent might have been, given various facts about her history, are limited to reveries about what might have happened to shake the agent loose from the bonds of the manipulator. When, on the other hand, an agent's pattern of response to reasons is a result of neutral causal forces, imagining the life she might have had requires imagining less drastic changes from her actual history: we need imagine only luck that overrode neutral misfortune distributed at random and happening to land on the agent, rather than, as in the other case, imagining luck that overrode the determination of a manipulator to inflict misfortune on the agent. This is why more powerful manipulators undermine their victims' freedom more thoroughly. The more powerful the manipulator, the less that could have happened to intervene to create an agent who responds to reasons in some way other than the way in which the manipulator designs.[10]

---

[10] Modal arguments such as that sketched in the last few paragraphs have pitfalls, so it is worth presenting the argument with greater precision. Consider, again, the Manipulated and the Unlucky: the Manipulated is a victim of indoctrination, the Unlucky has the same tendencies of response to reasons as the Manipulated as a result of neutral causal forces. Call the person who manipulates the Manipulated "the Indoctrinator", and call the set of features that cause the Unlucky to have the choice-producing mechanism she has "the Curse". Let W1 be the set of all possible worlds in which two things are true: (1) the Indoctrinator is aiming to produce just the choice-producing mechanism in the Manipulated that the Manipulated actually has, and (2) the Manipulated has a different choice-producing mechanism. The members of W1, then, are possible worlds in which something interferes to prevent the Indoctrinator from accomplishing his aims. Let W2 be the set of all possible worlds in which two things are true: (1) the Curse is present, and (2) the Unlucky has a different choice-producing mechanism from his actual choice-producing mechanism. The members of W2 are possible worlds in which something interferes to prevent the Curse from having its effect. Let R be a mapping from possible worlds onto other possible worlds as follows: For all possible worlds x and y, xRy iff x

There are some consequences of the view just sketched that are worth highlighting. We can imagine neutral causal forces that happen to track the formation of just the same patterns of response to reasons as manipulators track. This is the idea, in fact, of the robotic tutor: it teaches your children just as a living, breathing tutor would—while it doesn't actually care about what kind of responses to reasons your children evince (being a machine, it doesn't care about anything at all), it produces in them just the same mechanisms of response that a person who did care would produce. And, most importantly, the robotic tutor, like the real tutor, tracks this result. Conversely, we can imagine manipulators who are no better at producing in their victims the desired tendencies of response to reasons than various neutral causal forces would be. That is, manipulators can be weaklings, or they can have a failure of nerve. The monarch might imprison the Kasper Hauser figure and instruct his servants not to do anything at all to prevent rescue of the child, or the monarch might be determined to prevent rescue at all costs and simply lack the control over his servants needed to see to it that attempted rescues would be prevented. In cases such as these, either neutral causal forces that actually produce a particular pattern of response to reasons on the part of some agent also track that response, or else a manipulator who, in fact, brings about a particular mechanism of response to reasons in some agent fails to track the production of that mechanism. Are these counterexamples to the offered explanation for the freedom-undermining force of manipulation? I think not. That is, I am willing to accept the implied consequences: neutral causal forces that function like robotic tutors do undermine the freedom of their victims, and manipulators that function no more effectively than non-tracking neutral causal forces do not.[11]

---

contains the Indoctrinator, and y contains the Curse, and all other differences are deductive and causal results of this difference; y results, that is, from replacing the Indoctrinator with the Curse in x. Now, consider possible worlds, y1 and y2: in both an Obstacle is present that is such as to defeat the actual causal effect of the Curse, but would be circumvented by the Indoctrinator, and y1Ry2. Since the Indoctrinator tracks the production of the Manipulated's actual choice-producing mechanism, and the Curse does not track the Unlucky's actual choice-producing mechanism, there must be some such Obstacle. y2 is a member of W2, but y1 is not a member of W1. Further, there are no possible worlds x1 and x2 such that x1Rx2 and x1 is a member of W1, but x2 is not a member of W2: if there were such a pair of worlds, then there would be some force that would prevent the causal impact of the Indoctrinator but not prevent the causal impact of the Curse; but the Curse is, by definition, a force that produces the actual, but not the counterfactual effects produced by the Indoctrinator. It follows that for all w1, members of W1, there is one and only one member of W2, w2, such that w1Rw2; however, there is at least one w2, a member of W2, such that there is no w1, a member of W1, such that w1Rw2. What this implies is that W1 is a "smaller" set than W2. And it follows that there are fewer choice-producing mechanisms possible for the Manipulated than there are for the Unlucky.

11 It might be objected that the consequence of the proposed view accepted here is unsatisfactory (and, hence, the proposal is unsatisfactory) since an agent's freedom is not

These consequences are not so implausible. In the case of the robotic tutor, there is a strong draw towards believing that the robotic tutor is merely the tool of a manipulator who constructed the robot merely as a means for efficiently manipulating the victim. This needn't be what actually happens—perhaps the robotic tutor rises by chance from a pile of microchips that happened to have been sitting next to the perfectly placed power-source—but the draw towards imagining such a hand behind the construction of the robot comes from the recognition that the robot is functionally no different from a manipulator, and so appears to be the extension of a manipulator. That is, we recognize, however unconsciously, that the danger to freedom comes from being in the hands of a thing that tracks the production in you of a particular kind of tendency to respond to reasons. While such tracking is often associated with the presence in the tracker of certain desires or intentions, what matters for freedom is the tracking, not the desires or intentions.

Similarly, while Kasper Hauser's freedom might be damaged by the monarch even when the monarch fails to track the production in him of the tendencies for response to reasons that his treatment actually produces, the damage is not the kind of damage that is distinctive of manipulation. Such a Kasper Hauser may lack certain capacities that are required for freedom—he might not know, for instance, how to do various things that he would need to know how to do were he to be free to do those things—but his freedom is then undermined not because he is manipulated, *per se*, but because of mis-

---

undermined when she is caused to respond to reasons in a particular way unless the cause has some sort of ill will towards her. But this objection conflates being damaged with having a right to restitution. When a force such as a robotic tutor that neither has, nor can have, either good or ill will towards us nonetheless produces and tracks the production in us of a particular pattern of response to reasons, we are damaged, but we are not owed restitution, since the robotic tutor is not the kind of thing from which restitution can be rightly demanded. On the other hand, when a person plays this same role in producing our tendencies of response to reasons we are not just damaged but we are also owed restitution by that person. This point, however, is entirely compatible with the explanation for the freedom-undermining force of indoctrination offered here.

There is another way of motivating this same kind of objection. It might be suggested that we are unfree as a result of indoctrination because we are being objectified, or being treated as a means, rather than an end-in-itself, by the indoctrinator. Since a neutral causal mechanism, like a robotic tutor, can't have good or ill will towards us, it can't treat us with or without the kind of respect that our autonomy demands. The appeal of this line of thought comes, I believe, from the thought that there is some sort of constitutive connection between freedom (or autonomy, anyway) and being treated as a member of a moral community. What indoctrinators do to us, on this line of thought, is to exclude us from the moral community by failing to show respect for our autonomy. But this form of the objection requires defense of a claim that is inordinately difficult to defend. Namely, that you can't be free if you aren't treated by others as though you are. This claim has the unacceptable consequence that attempts at manipulation undermine freedom even when they fail, since, after all, attempts at manipulation also involve treating others as though they are not free.

fortune he happened to suffer as a result of manipulation that not all acts of manipulation would produce. For him to lose freedom for the distinctive reasons that manipulation diminishes freedom, the forces acting on him must track his pattern of response to reasons. They must, that is, limit the range of possible lives he might have led.

So, indoctrination undermines freedom by providing stricter limits on the agent's pattern of response to reasons than are placed on the agent by neutral causal forces that cause her to respond to reasons as she does. Notice that this is also true of agents who come to respond to reasons perfectly as a result of indoctrination. Intuitions are likely to diverge on the question of whether this is a welcome or unwelcome result. For those who hold that indoctrination undermines freedom of will even when it improves its victim, nothing more needs to be said. An explanation has been given for that intuition: a great deal would have had to change to divert her from goodness, to make her, say, the devil's child, and thus that many fewer ways of responding to reasons were available to her. However, there is a case to be made for the opposite view under which agents who have been induced to respond to reasons in some way superior to the way in which they would have responded in the absence of manipulation are actually *liberated* by the influence of the manipulator. After all, we might say, such "indoctrination" is better described as "education".

But the above explanation for the freedom-undermining effect of indoctrination can be maintained even by those who accept this line of thought. To accept this line of thought is to see freedom of will as constituted by correct responsiveness to value.[12] It is to insist that a causal influence on one's choices takes away from one's freedom only if there is something bad about having one's choices so influenced. If the relation between a manipulator and a victim can be duplicated without duplication of the evaluatively negative features usually exemplified by such relations, then, on this line of thought, someone can be manipulated without loss of freedom of will. But notice that this is entirely consistent with the explanation for the freedom-undermining effect given above. That explanation does not require the claim that the mere fact that fewer ways of responding to reasons are available to an agent is sufficient for her freedom to be diminished. Rather, what is claimed is that when we ask what kinds of lives were available to such agents we find that there were fewer available to the indoctrinated than to those who have come to be as they are through neutral causal forces. If we think that this question does

---

[12] Such a position is offered in Susan Wolf, *Freedom Within Reason*, Oxford University Press, Oxford, 1990. See my discussion of what I call "self-transcendence" views of freedom of will in "Free Will and Agency at It's Best" in *Philosophical Perspectives* v. 14, James Tomberlin ed., Oxford University Press, Oxford, 2000. As I argue in *Liberty Worth the Name: Locke on Free Agency*, Princeton University Press, Princeton, 2000, Locke holds a related view.

not apply to the agent before us, for some reason or another, then we are not likely to see her manipulation as freedom-undermining even if fewer ways of responding to reasons are available to her as a result of manipulation. But this is precisely what someone who sees improvement through manipulation as having no damaging effect on an agent's freedom holds. For someone who thinks of freedom this way, of agents improved by manipulation the question of unfreedom never arises, for they have been given a gift by the manipulator rather than twisted into something imperfect. We might add: you have nothing worthwhile taken from you when you are made to respond to reasons in some way superior to the way in which you previously responded, and so there is no relevant question about what kind of agent you might have been. We care about that question because we care about how you might have been *better*; if you can't have been better, there's no question to ask.

The various modal notions invoked in the explanation for the freedom-undermining force of indoctrination—such as the notion of more lives being *available* to the agent, or there being less that *could* have intervened to alter the agent's situation—are extremely weak. In fact, for all that has been said, they are weaker even than the notion of "could" offered by those who subscribe to the conditional analysis of "could have done otherwise". And they are that much weaker yet than the kinds of modal notions that incompatibilists employ and argue to be the only sort relevant to freedom. In fact, no more is being brought in through the usage of these modal notions than merely the assertion of the existence of a possible world with certain features. All that is being claimed is that, relative to any non-deviant similarity metric that one should choose, there are more possible worlds in which the Unlucky responds differently to reasons than possible worlds in which the Manipulated does. Incompatibilists insist that should determinism be true, the existence of possible worlds in which agents do various things other than what they do in the actual world, or have various properties they lack in the actual world, are utterly irrelevant to freedom. After all, incompatibilists insist, those worlds are not accessible to the agent; the agent cannot get to them from here. However, the explanation offered above for the freedom-undermining force of indoctrination can be accepted even by an incompatibilist who insists that the difference between the Manipulated and the Unlucky isolated here is of no relevance whatsoever to freedom. For such an incompatibilist, the explanation should be taken as a diagnosis of a non-veridical intuition. I shall explain.

Incompatibilists neutral on whether or not determinism is true are able to explain the belief in our freedom—and the notable fact that it doesn't go away in those who accept determinism—by pointing to our ignorance or our irrationality. In its most common form this explanation involves insisting that if determinism is true, the reason we think we can do otherwise derives from

our ignorance of the impact of the past and the laws on our choices, or if we do know what there is to know about such things, then belief in our freedom comes from an irrational failure of application of that knowledge to our own case. But the incompatibilist could explain our resolute belief in our freedom in the face of determinism in another way. The incompatibilist might suggest that our belief in our freedom comes from our failure to disambiguate various compatibilist notions of "can" and "could" from the sort that are really of importance to freedom. Perhaps, that is, we would continue to believe ourselves to be free even after, for instance, calculating what we will do tomorrow from consideration of the past and the laws because we make a philosophical error: we mistakenly think that, for instance, the conditional analysis of "can" captures the sense of "can" that is relevant for freedom. An incompatibilist who takes this route sees the mistaken belief in our freedom as deriving from the pernicious acceptance of the theory that the mere existence of a possible world with certain features in which the agent chooses differently is enough for freedom to so choose.

The explanation for the freedom-undermining force of indoctrination offered here can be accepted by an incompatibilist by making a very similar move. Such an incompatibilist can say that the source of the mistaken belief that the Manipulated is less free than the Unlucky derives from mistaking genuine availability of alternatives for the weak kind of availability invoked in the explanation and then coupling this confusion with the recognition that fewer lives are available, in this weak sense, to those who are indoctrinated. This leads us, the incompatibilist can say, to the mistaken intuition that indoctrination undermines freedom where neutral causal forces do not. That is, the incompatibilist can use the explanation offered here to diagnose the intuition without asserting that the intuition is veridical. I myself do not accept this way of construing the explanation. It seems to me much more likely that the existence of certain possible worlds is all that is required for freedom; the difficult question is what those possible worlds need to be like to be of relevance to freedom, and what kind of similarity metric needs to be used in determining which are relevant to freedom and which not. But nothing about the explanation for the freedom-undermining force of indoctrination given here depends upon the truth of these very controversial claims (for which no argument has been provided here).

Coercion

It is not immediately obvious how to extend the strategy of explanation sketched in the previous section to the case of coercion, but, as I will argue in this section, the strategy can be extended. Recall the distinction between coercion and indoctrination: where freedom-undermining indoctrination produces in its victim a particular mechanism for response to reasons, coercers,

rather, capitalize on the mechanisms for responding to reasons that their victims already possess. Coercers, that is, give their victims reasons for acting as the coercers want them to act. In standard cases—such as that of the cashier who gives in to the gun-wielding assailant—the coercer takes advantage of the fact that the victim responds to reasons appropriately: the coercer gives the victim, for instance, genuine self-interested reasons for acting as he wants her to act. This, however, is not necessary. The coercer may be aware of the victim's imperfect tendencies to respond to reasons and may manipulate the victim only by bringing it about that she *takes herself* to have reason to do what he wants her to do, even when she does not, in fact, have any reason to do it. A hydrophobe might be effectively manipulated to choose to do something deeply self-destructive, for instance, through threats of being thrown into open water, even if the damage suffered through the compliant choice would be far worse than the damage that could be rationally expected to result from being thrown into the open water (even taking into account the pain suffered by the hydrophobe given his aversion to water). This distinction, however, plays no role in the explanation for the freedom-undermining force of coercion: the explanation to be given here should explain both the freedom-undermining force of coercion that supplies actual reasons for compliance and coercion that supplies only the appearance of reasons to its victim.

The key to the explanation for the freedom-undermining force of coercion is that, as a general rule, coercers don't merely produce, but also track, the compliance of their victims. A robber who threatens to injure the cashier should the cashier not hand over the money would usually be ready to threaten a more serious injury were the cashier to prove unresponsive to the first threat. That is, the coercer is rarely attached to the particular nasty consequence that he threatens; with some limits, he is ready to bring about whatever consequence would serve to bring the victim around to compliance. What this implies is that there are a fairly wide range of possible patterns for response to reasons all of which will result, given the coercer's commitment, in the victim's compliance. If what the victim cares about is the avoidance of pain, the coercer will threaten pain; if what the victim cares about is the avoidance of death, the coercer will threaten to kill; if what the victim cares about is the welfare of her family, the coercer will threaten her family. The coercer tailors his threat to the features of the mechanism for response to reasons that the victim possesses.

Coercers, of course, do not do this to an unlimited degree. Some threats are simply beyond the pale. Parents, for instance, are often willing to threaten early curfews even when they are entirely unwilling to threaten electric shocks. But, still, within certain ranges, coercers do track their victims' compliance through a willingness to make different threats should the actual ones be ineffective. Coercers simply vary in their degree of resolve.

Sometimes, as in the case of indoctrination, we care about having available to us a range of possible patterns of responsiveness to reasons and we seem to have a more limited range available to us given the fact that a manipulator is tracking one particular sort of pattern (usually, the sort that serves the manipulator's ends). However, we also care about what choices we make through use of the particular mechanisms of response to reasons with which we are invested. That is, sometimes we count differences between mechanisms by counting differences in sets of possible worlds associated with each mechanism. That is, we alter the circumstances and compare the patterns of response across the resulting possible worlds that is entailed by the possession of one mechanism with the pattern entailed by the possession of another; when the patterns are different, the mechanisms are different. Call this the "pattern based" criterion of individuation of choice-producing mechanisms. But there is another, second, criterion: Sometimes we take two mechanisms to be similar enough to be thought the same just when each issues in the same choice in our actual circumstances. In this second sense, we hold fixed the actual circumstances and vary the mechanisms; when two mechanisms issue in the same choice in the actual circumstances, they are the same mechanism. Call this the "function based" criterion of individuation of choice-producing mechanisms; under this criterion two choice-producing mechanisms are the same when they are functionally the same; when, given the circumstances the agent faces, changing from one mechanism to the other doesn't change the agent's choice.

So, for instance, we will spot an important difference in the pattern-based sense, and not in the function-based sense, between an agent who chooses to issue a hurtful remark to a friend and would not have done so had she expected her friend to break down crying as a result, and someone who issues the same hurtful remark and would have chosen to do the same thing even if she had expected the tearful response. However, in the function-based sense, these two choice-producing mechanisms are really no different: given the actual facts, in which there was no reason to expect the friend to take things so hard, both mechanisms result in the same choice. Depending on whether or not we are concerned primarily with the agent's enduring qualities or primarily with the particular choices she made, we will either see the pattern-based or the function-based way of individuating choice-producing mechanisms as the crucial criterion of individuation. Sometimes, that is, the friend will be mollified to be told, "I wouldn't have said that if I'd thought it would make you cry", and sometimes the appropriate response to this excuse on the part of the friend is just, "But you still said it, didn't you?"

Whether more or fewer forms of response to reasons are available to us depends on how we individuate different mechanisms for responding to reasons. But, in fact, I believe that when we worry about freedom we are

interested in having a wider, rather than a narrower, range of mechanisms available to us in both the sense of "wide" and "narrow" associated with the pattern-based criterion for individuation, and the sense of "wide" and "narrow" associated with the function-based criterion. Indoctrination narrows the different ways in which we might live in the first sense: it narrows the number of different patterns of response to reasons that are possible for us. Coercion, on the other hand, narrows the different ways in which we might live in the second sense: coercers make the differences that are possible for us functionally equivalent. They see to it that we will choose as they want us to choose across a wide range of differences in our dispositions for spotting and responding to reasons. Coercers do this by making irrelevant the differences between, say, someone who places the safety of her children above her own safety and someone who does not. The coercer will see to it that either person gives in to his demands by issuing the threat that is appropriate to the pattern of response to reasons exhibited by the agent. What this means is that when we are coerced, fewer functionally different mechanisms for responding to reasons are possible for us.[13]

---

[13] It is worth pausing, once again, to put the modal argument just sketched into a more precise form. Analogous to the Manipulated, the Unlucky, the Indoctrinator and the Curse of the previous section, consider the Coerced, the Unfortunate, the Coercer, and the Misfortune: The Coerced is the actual agent who suffers coercion at the hands of the Coercer; the Unfortunate has exactly the same psychological dispositions for recognizing and responding to reasons as the Coerced, but finds herself having to make and making just the choice that the Coerced makes only as a result of the Misfortune, the presence of which results in the Unfortunate having just as much reason to choose as the Coerced chooses; but unlike the Coercer, the Misfortune does not track that choice. Let W1 be the set of all possible worlds minimally different from the actual world but in which (1) the Coerced chooses otherwise than she actually chooses, and (2) the Coercer is aware of the kinds of reasons that the Coerced responds to, and makes an effort to provide her more reason to choose as she actually chooses than to choose in any other way. The members of W1, then, are possible worlds in which the Coercer tries and fails to induce compliance from the Coerced. Let W2 be the set of all possible worlds minimally different from the actual world but in which (1) the Unfortunate chooses otherwise than she actually chooses, and (2) the Misfortune functions just as it actually functions. As before, let R be a mapping of possible worlds to possible worlds where xRy just in case y results from replacing the Coercer in x with the Misfortune. Since the Coercer tracks the compliance of the Coerced and the Misfortune does not track the compliance of the Unfortunate, there must be some choice-producing mechanism, M, such that there are possible worlds x1 and y1, such that the Coerced has M in x1 and chooses just as she actually chooses, and the Unfortunate has M in y1 and chooses otherwise than she actually chooses, and x1Ry1. This must be so, since what the Coercer does, and what the Misfortune does not, is to see to it that some choice producing mechanisms that are different in the "pattern-based" sense are the same in the "functional" sense. x1 is not a member of W1 and y1 is a member of W2. It follows that for all w1, members of W1, there is one and only one member of W2, w2, such that w1Rw2; however, there is at least one w2, a member of W2, such that there is no w1, a member of W1, such that w1Rw2. As before, what this implies is that W1 is a "smaller" set than W2. And it follows that there are more functionally different choice-producing mechanisms available to the Unfortunate than there are available to the Coerced.

There are a few consequences of the argument just sketched that are worth pointing out. First, notice that those who issue offers to others, rather than threats, seem to function just as coercers do: they induce compliant choices by creating reasons for their "victims" to choose as they want them to choose; and benefactors of this kind usually track the compliant choice in just the same sense in which coercers track compliant choices: they are willing to tailor their offers to the particular forms of responsiveness to reasons possessed by the recipients of their offers.[14] If I am committed to bringing it about that you choose to come work for me, and I believe you to care about money, I offer you the amount you want. If you care about freedom, I offer you more vacation time; if you care about moving expenses, I offer to pay them. There are, of course, things that I won't offer you, no matter how much you want them, but, still, to some degree at least, I track your compliance just as coercers track their victim's compliance. But, intuitions will vary on the question of whether or not compelling offers undermine freedom of will. Can the explanation for the freedom-undermining effect of coercion be accepted by those who think they do not? The answer is "yes".

Recall that in the case of indoctrination, indoctrinators who induced in others better mechanisms for responding to reasons could be consistently thought not to undermine their freedom, since they conferred on them nothing but a benefit and thus made the question of what alternative lives were available to such agents moot. Similarly, those who hold that offers do not undermine freedom, even though threats do, can say that benefactors, as opposed to coercers, confer on others nothing but a benefit and thus the question of their impact on freedom is moot. Or, put more precisely, benefactors limit the number of functionally different choice-producing mechanisms available to agents only by making all of the functionally different mechanisms—all of the mechanisms that lead to compliant choice—better than they would have been in the absence of the determination on the benefactor's part to induce compliance in the recipient.[15] For those who think of freedom

---

[14] It is actually quite difficult to provide a good test for distinguishing threats from offers. See Robert Nozick, "Coercion" in *Philosophy, Science, Method: Essays in Honor of Ernest Nagel*, S. Morgenbesser, P. Suppes, and M. White, eds. St. Martin's Press, New York, 1969, pp. 440-472. And, Harry Frankfurt, "Coercion and Moral Responsibility" in *The Importance of What We Care About*, Cambridge University Press, Cambridge, 1988, pp. 11-25.

[15] This is why, for instance, some who think that offers are not in general freedom-undermining might still consider certain *exploitative* offers to be freedom-undermining. In particular, if I know that you are desperate to please me and I offer you a pat on the head if you will lick my boot, and you do it, I might be thought, even by those who take many offers to have no detrimental effect on freedom, to be undermining your freedom. However, in this case, my offer hasn't turned your tendency to respond to anything that will get my approval as a reason for action into a tendency worth having; quite the reverse. The thought that my offer is exploitative and freedom-undermining comes from the

of will as essentially worth having, the question of unfreedom never arises in the case of those who issue offers, but only in the case of those who issue threats. Such theorists can accept the explanation for the freedom-undermining effect of coercion offered here by saying that we care about having multiple functionally different choice-producing mechanisms available to us only because we want to have better choice-producing mechanisms available. When another person provides us with better mechanisms than we could have had without her, the question of whether more or fewer are available to us becomes moot.[16]

As in the case of indoctrination, one of the results of this discussion is that an agent's freedom can be diminished even when the coercer is replaced with a neutral causal mechanism that happens to track the agent's compliance in just the way that a coercer usually tracks her compliance; and, also, cases of coercion that do not involve tracking the compliance of the victim in the way described do not diminish the victim's freedom. As before, these results are not objectionable, but welcome. There really is no difference between throwing cargo overboard as a result of a storm, and throwing it overboard to mollify pirates unwilling to make a more effective threat should the captain refuse to comply. The threat to our freedom that coercion poses does not come from the fact that we face awful consequences for non-compliance—storms, for instance, place the same constraints upon us. The threat to our freedom comes from the readiness on the part of the coercer to see to it that we won't brave the consequences of non-compliance; that is, the threat comes from the fact that coercers are willing, to some degree at least, to see to it that we comply with their demands. To put the point in the context of the present example: the storm doesn't care if we hold the cargo and allow for the sinking of the ship rather than throw the cargo overboard, while, standardly at least, pirates prefer the latter option to the former. Still, pirates that really don't track the satisfaction of this preference eliminate our freedom no more than storms do. As in the case of indoctrination, it is the tracking that matters, not the sentience of the force doing the tracking.

Conclusion

This paper began with the question of why it is that our freedom seems to be undermined when other agents see to it that we make certain choices even though it seems that the same causal pressures issued by a non-agent only face us with misfortunes of the sort that agents face all the time without any limitation placed on their freedom. The answer, in short, is that agents track

---

thought the pat on the head wasn't worth the degradation of licking my boot, no matter what you happened to think.

16 Relevant here is Gerald Dworkin, "Is More Choice Better than Less?" in *Midwest Studies in Philosophy*. v. 7, 1982, pp. 47-61.

the compliance of their victims and neutral causal forces do not. Notice, however, that the kind of tracking of compliance that coercers and indoctrinators exhibit involves capacities on their parts for recognizing and responding to reasons. Coercers and indoctrinators track the features of the world that they track by virtue of the relationship that those features bear to reasons. Indoctrinators track certain choice-producing mechanisms that are identified by virtue of the pattern of response to reasons exhibited by agents possessing those mechanisms; coercers induce compliance on the part of their victims by recognizing what would count for the victim as a reason to act. Coercers and indoctrinators, then, usually exercise their own capacities for the recognition of reasons; in fact, it is in part because they have such capacities that they are capable of tracking their victims' compliance in the way in which they do.

In so far as the kind of tracking of response exhibited by a coercer or an indoctrinator normally involves the use of capacities for recognizing and responding to reasons, it also normally involves exercise of will on the part of the coercer or indoctrinator. Our wills function as they do in part through the exercise of such capacities. But, the tracking of features of the world picked out by virtue of their relation to reasons is not sufficient for the possession of a will; the robotic tutor doesn't have a will, because it doesn't have a mind, but it tracks the production of tendencies to recognize and respond to reasons. Still, it is a fact that this kind of tracking almost always goes hand in hand with the possession of a will on the part of the entity doing the tracking. Given this, it is really no surprise that we tend to see the causal role of the will of another as having a peculiar kind of significance in the limitation of our freedom. We confuse, that is, the feature by virtue of which the entity applying pressure on us limits our freedom (the fact that it tracks compliance) with what it is about that entity that *gives it* the feature by virtue of which it limits our freedom (the fact that it has a will). But this point does not blunt the primary point for which this paper has argued: to be in the hands of another is usually importantly different from being in the hands of Nature; and this difference makes—or, at least, quite naturally seems to make—a difference to freedom of will.[17]

[17] Thanks to Michael Bratman, John Fischer, Alfred Mele, Elijah Millgram, Jennifer Rosner and two anonymous referees for very helpful comments and suggestions.