PREVENTION, COERCION, AND TWO CONCEPTS OF NEGATIVE LIBERTY

Michael Garnett

In *Without Trimmings: The Legal, Moral, and Political Philosophy of Matthew Kramer* (Oxford University Press, 2022), pp. 223-38

According to Thomas Hobbes, 'Liberty, or FREEDOME signifieth (properly) the absence of… externall Impediments of motion'.[1] On this pure, negative view, liberty requires only the absence of physical prevention. It's a conception that's been ingeniously defended and developed in recent years by philosophers such as Hillel Steiner, Ian Carter and Matthew Kramer, all of whom argue for sophisticated versions of the view that one's degree of freedom is a function of the range of bodily movements that one is physically unprevented from performing.[2] We find the core of this traditional view succinctly expressed, for instance, in Kramer's *F Postulate*: 'A person is free to φ if and only if he is unprevented from φ-ing'.[3]

However, not every influential figure in the negative tradition chose to characterise freedom in terms of *nonprevention*. Others did so, instead or as well, in terms of *noncoercion*. Jeremy Bentham, for instance, tells us that 'Liberty… is neither more nor less than the absence of coercion'.[4] Similarly, in his 'Two Concepts of Liberty' Isaiah Berlin equates 'the "negative"

---

[1] *Leviathan*, ed. Richard Tuck (Cambridge: Cambridge University Press, 1991), 145.

[2] Hillel Steiner, 'Individual Liberty', *Proceedings of the Aristotelian Society* 75 (1974): 33–50; Ian Carter, *A Measure of Freedom* (Oxford: Oxford University Press, 1999); Matthew H. Kramer, *The Quality of Freedom* (Oxford: Oxford University Press, 2003).

[3] *The Quality of Freedom*, 3.

[4] UC: lxix.44. Quoted in Frederick Rosen, *Classical Utilitarianism from Hume to Mill* (London: Routledge, 2003), 247. This view was also repeated almost *verbatim* by the Eighteenth Century pamphleteer John Lind, according

conception of liberty in its classical form' with 'non-interference, which is the opposite of coercion'.[5] And in the opening pages of *The Constitution of Liberty*, F. A. Hayek defined 'a state of liberty or freedom' as 'that condition of men by which coercion of some by others is reduced as much as possible in society'.[6]

How important is this difference? The distinction between freedom-as-nonprevention and freedom-as-noncoercion isn't one that's received much attention. This is probably for the following reason. On the one hand, it's common for nonprevention views of freedom to restrict the scope of freedom-limiting preventions to those *brought about by other people*, on the grounds that freedom (in the politically-relevant sense) is an inherently social concept.[7] On the other hand, accounts of coercion often take coercion necessarily to involve at least some physical restriction of the victim's available options.[8] Putting these together, it's natural to think that both prevention (in the freedom-relevant sense) and coercion essentially involve

---

to whom liberty is 'nothing more nor less than the absence of coercion' (*Three Letters to Dr Price* (London: T. Payne, 1776), 16).

[5] 'Two Concepts of Liberty', in *Four Essays on Liberty* (London: Oxford University Press, 1969), 128.

[6] *The Constitution of Liberty* (Chicago: University of Chicago Press, 2011), 57. Note, too, that Hobbes himself admits a secondary sense of 'liberty' understood as the absence of coercively enforced rules (i.e., 'the Silence of the Law'): 'In cases where the sovereign has prescribed no rule, there the Subject hath the Liberty to do, or forbear, according to his own discretion' (*Leviathan*, 152).

[7] See, e.g., Berlin, 'Two Concepts of Liberty'; David Miller, 'Constraints on Freedom', *Ethics* 94, no. 1 (1983): 66–86; Kristjan Kristjánsson, *Social Freedom: The Responsibility View* (Cambridge England ; New York: Cambridge University Press, 1996); Steiner, 'Individual Liberty'.

[8] Indeed, this was Bentham's view, inasmuch as he characterised liberty as the absence of coercion, and 'the evil of coercion' as the evil of 'restraint' (*An Introduction to the Principles of Morals and Legislation*, 175). See also Christine Swanton, 'Robert Stevens on Offers', *Australasian Journal of Philosophy* 67, no. 4 (1989): 475; David Zimmerman, 'Coercive Wage Offers', *Philosophy and Public Affairs* 10, no. 2 (1981): 134; Alan Wertheimer and Franklin G Miller, 'Payment for Research Participation: A Coercive Offer?', *Journal of Medical Ethics* 34 (2008): 390.

being stopped from doing things by other people, and that the differences between them in this context are therefore negligible.

In this paper I argue that this obscures an important and useful distinction. Nonprevention and noncoercion are distinct goods. Although they often coincide, neither entirely suffices for the other. Moreover, negative theorists have reason to care about both of them. I therefore argue against the recent Hobbesian tendency to theorise negative freedom as nonprevention alone, arguing instead that a fully defensible negative account must draw on both of the tradition's core ideas: freedom-as-nonprevention and freedom-as-noncoercion.

1. Actions and Persons

1.1. A Distinction

I'm concerned with the distinction between prevention and coercion. In order to get it clearly in view, however, I need to begin with another. This is the distinction between the *freedom of a person* and the *freedom of an act*.

Sometimes we ascribe freedom and unfreedom to persons, or to groups of persons. We say that someone is or isn't a free person, or that they're more or less free now than they were before, or that they're free or unfree in certain respects or to do certain things. Other times, however, we ascribe freedom and unfreedom to actions or to choices. We say that someone acted freely or unfreely, or that they did or didn't make a free choice, or that they were or weren't free in performing some action.

Note that this distinction is different from that between *overall* and *specific* freedom. This latter distinction, frequently drawn in discussions of negative freedom, concerns the difference between judgements of the form 'some person is free' (or 'some person is freer than some other person'), on the one hand, and judgements of the form 'some person is free to

perform some action', on the other.[9] Both of these relate to the freedom of persons: in assessing the freedom of a person, one may assess their freedom *in toto*, or in specific domains, or with respect to certain actions. In each case, however, one is assessing the *person's* freedom to perform some action, and not the freedom of the *action* that they thereby perform. Hence the distinction between free persons and free actions precedes the distinction between overall and specific freedom.

To illustrate: suppose that a mugger threatens to stab me if I don't hand over my wallet. In this case, I'm free to hand over my wallet: it's a live option for me and nothing prevents me from doing it. Therefore, I have the specific freedom of handing over my wallet. When I do hand it over, however, I don't hand it over freely. I hand my wallet to the mugger under coercion. Hence the action I perform is an unfree action. Nevertheless, it's an unfree action that I was free to perform. If we want to assess my degree of personal freedom, then, we need to know what range of actions I'm free to perform, and so we need to know (*inter alia*) whether I'm free to perform the specific action of handing my wallet over to the mugger. If we want to assess the degree to which I'm acting freely, by contrast, we need to look at the circumstances of the action that I actually perform. Thus specific freedom is not freedom predicated of actions, but freedom predicated of persons with respect to their actions.

Judgements about the freedom of actions are often motivated by concerns about moral responsibility. For this reason, many accounts of free action are attempts to specify the conditions that must be met for an act to be morally attributable to an agent. However, our interest in free action isn't exhausted by concerns about moral responsibility. There's a distinctive good associated with acting freely, and some accounts of free action are attempts to elucidate that good in its own right. There may be intrinsic value, for instance, in acting

[9] Carter, *A Measure of Freedom*, 12–14; Kramer, *The Quality of Freedom*, 9.

creatively or with a free spirit; attempts to elucidate this value are attempts to provide an account of free action, but not one that's of any direct relevance to questions about moral responsibility.[10] Similarly, it may be valuable in itself to act in ways uninfluenced by coercion or the illegitimate interferences of others. One might think, for instance, that a person with severe mental illness who doesn't meet the competence threshold for moral accountability still loses something of value when their actions are coerced. In elucidating such a value, again, one will be giving an account of free action that isn't (or, at least, isn't yet) one that's directly relevant to questions about moral responsibility.

Indeed, it's worth noting that exactly parallel points about the detachability of our interests in freedom and responsibility apply with respect to the freedom of persons too. Just as one might be interested in what it is for an *action* to be one for which a person can be held morally responsible, one might also be interested in what it is for a *person* or *agent* to be the sort of thing that can be held morally responsible. To be a morally responsible agent is, most likely, to be a free agent in some morally important sense. But this isn't the sense of personal freedom that theories of socio-political freedom typically seek to understand. Such theories target a different form of personal freedom, one that's valuable independently of any role it may play in constituting one's status as a morally responsible being. In the same way, we may seek an account of what it is for a person to act freely in a socio-political sense without concerning ourselves directly with the question of what it is for a person to act with moral responsibility. Hence freedom may be predicated of both persons and actions, and it may be taken to pick out distinctively socio-political values in both cases.

1.2. Free Actions and Free Persons

[10] See e.g. Jonathan Gingerich, 'Freedom's Spontaneity' (Dissertation, University of California, Los Angeles, 2018).

Coercion is a property of actual actions or bodily movements: one cannot be coerced into doing something that one does not, in fact, do.[11] Yet with prevention the opposite is the case: one cannot do that which one has been prevented from doing. Hence to be coerced is to *act* unfreely, whereas to be prevented is not to act at all, but to *be* unfree. That is, coercion relates primarily to the freedom of actions, whereas prevention relates primarily to the freedom of persons. In order to understand the relationship between coercion and prevention, then, it's worth taking a step back and thinking about the general conceptual relationship between the freedom of a person and the freedom of their actions. What logical connections, if any, obtain between the two?

Some theories posit a very tight connection. Positive freedom theorists, for instance, typically take free action as basic and analyse personal freedom entirely in terms of it. That is, they begin with an account of what it is to act with true freedom—such as that to act freely is to act rationally, or from a correct understanding of the world, or authentically, or in accordance with one's essential nature—and then take a free person to be someone who succeeds in acting like that. On most positive views, that is, a free person *just is* someone who acts freely. Hence for the Stoics, the free person just is one who acts in accordance with reason (i.e., freely); for Spinoza, the free person just is one who acts on the basis of adequate ideas (i.e., freely); for Marx, one is truly emancipated only when one acts from consciousness of oneself as a species-being (i.e., freely); and so on.[12] Crucially, none of these theorists take the mere *capacity* or

[11] Note the disjunct: if a police officer says 'freeze or I'll shoot!' and you freeze, your freezing is an action—something you do, in the relevant sense—despite being an absence of bodily movement.

[12] See e.g. John M. Cooper, 'Stoic Autonomy', in *Autonomy*, ed. E. F. Paul, F. Miller Jr., and J. Paul (Cambridge: Cambridge University Press, 2003), 1–29; Michael J. Kisner, *Spinoza on Human Freedom: Reason, Autonomy and the Good Life* (Cambridge: Cambridge University Press, 2011); Allen W. Wood, *Karl Marx* (Henley-on-Thames: Routledge, 1981). Note, again, that such theorists don't take these to be requirements for moral responsibility; if they did, then their positions would entail that hardly anyone is ever morally responsible for anything.

opportunity for free action to suffice for personal freedom: to be a free person, on their views, one must actually succeed in acting freely. Hence someone who had the option of acting from adequate ideas but somehow chose to act from inadequate ones instead would not, for Spinoza, qualify as a free person; someone who was able to act from their essential human nature but chose to engage in alienated activity instead would not, for Marx, qualify as a free person; and so on.

This is the point of Charles Taylor's famous comment that positive accounts treat freedom as an 'exercise concept'.[13] Negative accounts, by contrast, take freedom to be what he calls an 'opportunity concept', which is to say that they *don't* analyse the freedom of persons entirely in terms of the freedom of their actions. On a Hobbesian view, for instance, a free person is one who is physically unprevented from performing (some sufficiently wide range of) possible actions. There is no entailment here to free action: such a person needn't act freely since, apart from anything else, such a person needn't act at all. Indeed, we might imagine someone confronted by a range of options so wide that they become paralysed by indecision and fail to act; negative theorists will typically ascribe a high degree of freedom to such a person, regardless of the fact that they perform no actions and therefore no free actions.

One way to construe the fundamental difference between positive and negative accounts of freedom, then, is in terms of the contrary positions they take on the relationship between the freedom of persons and the freedom of their actions: negative views, unlike positive ones, allow the two to come apart. Moreover, since negative views allow them to come apart, it would seem that negative theorists owe separate accounts of them. And this, one might think, explains how the ideals of nonprevention and noncoercion might fit neatly together in a complete negative theory of freedom, with the former constituting the account of personal

[13] 'What's Wrong with Negative Liberty', in *The Idea of Freedom*, ed. A. Ryan (Oxford: Oxford University Press, 1979).

freedom and the latter the account of free action. In turn, this suggests that insofar as negative theorists like Steiner, Carter and Kramer are concerned with analysing the freedom of *persons* they may be right to focus exclusively on nonprevention. Unfortunately, however, the pieces of this conceptual puzzle don't fit together quite so cleanly.

That's because even negative accounts have reason to recognise some systematic connections between the freedom of persons and the freedom of their actions. Suppose, for instance, that an enslaved person is subject to pervasive coercion and hence acts unfreely much of the time. Any plausible negative account should make it possible to infer from this that such a person lacks a substantial degree of freedom. It would be an odd view of freedom—and, in particular, an odd interpretation of the idea of negative freedom—that allowed the idea of personal freedom to become so untethered from that of free action that such a person could qualify as a substantially free person. For this reason, negative accounts must, and typically do, accept that the freedom of persons necessarily correlates to some extent with (at least some aspects of) the freedom of their actions.

Moreover, recognising such conceptual connections needn't undermine the negative theorist's principled opposition to 'exercise'-based, positive accounts. This is for two reasons. First, taking personal freedom to require the absence of certain types of unfree action is different from taking it to require the presence of free action. Hence the negative theorist can continue to treat personal freedom (at a time) as consistent with inaction (at a time), and therefore as requiring no actual exercise of agency.

Second, in drawing some conceptual connections between the freedom of persons and the freedom of actions the negative theorist needn't accept the positive theorist's characterisation of the latter. That is, the negative theorist needn't agree that to act freely in a socio-political sense is to act rationally, or virtuously, or authentically. They can instead conceive free action in purely negative terms, such that to act freely is merely to act in the

absence of coercion or other illegitimate interference. Just as there are positive and negative conceptions of personal freedom, that is, there are positive and negative conceptions of free action—and in connecting one with the other, the negative theorist may adhere to negative versions of both.

Hence it's incumbent on a negative account of the freedom of persons to demonstrate that, insofar as people act unfreely in a relevant negative sense (for instance, under coercion), they thereby lack at least some degree of personal freedom. As I explain in the next section, this is something that contemporary negative freedom theorists have succeeded in demonstrating to a considerable extent; as I argue in the rest of the paper, however, it's something that can never be achieved in full. The result is that nonprevention and noncoercion must be recognised as distinct components of a negative account of personal freedom: to be free one must not only be unprevented from acting but also uncoerced in acting, where these amount to different things.

1.3. Threats and Conjunctive Freedoms

Suppose that a mugger threatens to stab you unless you hand over your wallet, and that you hand over your wallet so as not to get stabbed. This is a paradigm case of coercion. As I've just been arguing, one would expect any plausible account of freedom—and, in particular, any plausible *negative* account of freedom—to yield the result that you suffer at least some loss of liberty here.

Not all negative accounts have done so, however. Hobbes, for example, denied that there is any loss of freedom-as-nonprevention (in his terms, 'corporal liberty') in this kind of case. On his view, one loses such freedom only when one is physically prevented from doing something. Yet conditional threats don't appear to involve any form of physical prevention. Nothing physically prevents you from handing your wallet over to the mugger; more

importantly, moreover, nothing physically prevents you from *not* handing your wallet over—yes, you'll get stabbed, but you'll still have your wallet. What prevents you from handing your wallet over is your *fear* of getting stabbed, and this fear isn't an external impediment of motion, but an internal one. Thus, Hobbes tells us, 'Feare, and Liberty are consistent', and 'all actions which men do in Commonwealths, for *feare* of the law, are actions, which the doers had *liberty* to omit'.[14]

Modern negative theorists, by contrast, generally seek to avoid this result.[15] They accept that standard cases of conditional threats, like the mugging case, involve diminutions in personal freedom. Moreover, they seek to deliver this result without giving up the basic Hobbesian analysis of freedom as the absence of physical obstacles to possible bodily movements. Thus they argue that nonprevention views of freedom needn't be committed to counterintuitive results regarding coercion cases, and that Hobbes was, in effect, overly hasty in deriving his conclusions on this matter. Key to these modern arguments is the idea of *conjunctive freedoms*.

Suppose that I'm free to go to the park today, inasmuch as nothing prevents me from going. Suppose that I'm also free to go to the beach today, inasmuch as nothing prevents me from going. Am I free to go to the park *and* to the beach today? Well, that depends. If the park and the beach are in opposite directions and both far from my house, then it may be physically impossible for me to visit both in a day. This third freedom, the freedom to go to the park *and* to the beach today, is a conjunctive freedom, and it's enjoyment isn't entailed by enjoyment of the two atomic freedoms that compose it. In assessing a person's overall freedom, then, we need to take account of their conjunctive freedoms, and this is what's missing from Hobbes'

[14] *Leviathan*, 146.

[15] An important exception is Steiner, who defends the austere Hobbesian view that 'neither the making of threats nor that of offers constitutes a diminution of individual liberty' ('Individual Liberty', 43).

analysis of conditional threats. When the mugger threatens to stab you unless you hand your wallet over, you do indeed retain the specific freedom of keeping your wallet. However, you lose the conjunctive freedom of keeping your wallet *and not getting stabbed* (or, more precisely, of keeping your wallet and doing any of the many things that you wouldn't be able to do had you just been stabbed). So the threat does physically prevent you from doing something—in fact, a whole range of things—and physical nonprevention views of freedom are therefore able to deliver the intuitive conclusion that coercive conditional threats diminish freedom.

This is the line of argument adopted by J. P. Day, Carter and Kramer, among others.[16] In this way, nonprevention views of freedom can be shown to yield plausible results concerning conditional threat cases. The impact of conditional threats on conjunctive freedoms means that, at least in a central range of cases, those who act under coercion have necessarily suffered some significant diminution in liberty. Nonprevention views therefore have the resources with which to establish plausible systematic connections between the freedom of persons and the freedom of their actions: it is no accident, on such views, that those who act unfreely (i.e. under coercion) are normally also to some extent unfree (i.e. prevented from acting in valuable ways).

Nevertheless, nonprevention views can't account for the freedom-limiting nature of coercion in its entirety. As I now argue, there are a number of cases in which prevention and coercion come apart, such that a person may be subject to active coercive control without being prevented from doing anything. These are cases of diminished personal freedom, and yet ones

[16] J. P. Day, 'Threats, Offers, Law, Opinion and Liberty', *American Philosophical Quarterly* 14, no. 4 (1977): 257–272; Carter, *A Measure of Freedom*, 224–32; Kramer, *The Quality of Freedom*, 194–210. As Kramer puts the point: 'the following pair of particular freedoms cannot be exercised conjunctively by the recipient of a genuine threat: the freedom to adopt the mode of conduct which the threatener is trying to discourage, and the freedom to undertake any mode of conduct that will be stymied by the threatener's infliction of a penalty for non-compliance' (195).

which nonprevention views struggle to handle. I focus on two such types of case: empty threats and preference manipulation. In each case I argue that there's a loss of negative freedom that can't be analysed in terms of prevention alone, and that contemporary negative theorists must therefore move away from prevention-only views and towards a view that recognises both prevention and coercion as distinct forms of personal unfreedom.

2. Coercion Without Prevention

2.1. Empty Threats

When someone issues a conditional threat, they normally make it *true* that the threat will come to pass if the victim fails to comply with their demands. Sometimes, however, they don't. This may be intentional, as when the threatener successfully bluffs, getting the victim to believe that the threat will be carried out when the threatener knows that it won't. It may also be unintentional. Consider:

> *Heart Chip: A* implants a chip in *B*'s heart that will kill *B* instantly if activated. *A* can activate the chip at any time by pressing a button on a remote control that *A* always carries. *A* uses this power to coerce *B* into serving *A*'s every whim. However, after five years the chip unexpectedly malfunctions and stops working. Nobody knows that this has happened. *A* and *B* continue as before, for a further five years.

In both of the two five-year periods *B* acts under credible threat of death. *A*'s effective coercive power over *B* is unaffected by the chip's invisible malfunction, and *B* continues to be subject to *A*'s will even after it stops working. Hence *B* acts unfreely for ten years: in stooping to shine *A*'s shoes in order to avoid instant death, that is, *B* isn't enjoying the good of free action, regardless of what's actually happening with the chip. Moreover, given that *B* acts unfreely

throughout both periods, there would seem to be a clear sense in which *B*'s freedom is reduced throughout both periods. Someone whose every act is coerced is, to some extent, an unfree person. Yet prevention-only views of freedom must hold that whereas *B* is significantly unfree for the first five years, *B* is then fully liberated by the chip's malfunction.

This is for the following reason. In the first five years, *A* has successfully prevented *B* from performing a wide range of sets of compossible actions that *B* would otherwise have been free to perform: specifically, any such set predicated on *B*'s defying *A* without being instantly killed. In the subsequent five years, however, *A* no longer prevents *B* from doing anything. In this subsequent period *B* is free to defy *A* without being instantly killed or (let us stipulate) suffering any other ill effects as a result. *B* is unaware of having these freedoms, but has them nonetheless. On prevention-based views, therefore, *B* suffers no unfreedom at the hands of *A* after the first five years. But this, I submit, is implausible: there is at least one clear sense in which *B*'s freedom remains diminished in the subsequent five years, on account of the coercion to which *B* continues to be subject; moreover, this is a negative, socio-political form of unfreedom that a negative account may fairly be expected to capture.

In making this argument, I don't mean to deny the value of nonprevention. The chip's malfunction really does free *B* to do a whole raft of things that *B* couldn't previously do, and it matters what a person can do even if they aren't aware of it. Theories of freedom need to account for this, and nonprevention views succeed in doing so. At the same time, however, we omit something crucial if we stop our analysis here. *B* doesn't *act* any more freely as a result of the chip's malfunction; *B* continues to act under coercion. And it would seem perverse to insist, as the prevention-only theorist seemingly must, that a person so pervasively subject to coercion can nevertheless be perfectly free.

Must the prevention-only theorist really hold that *B* is fully liberated by the chip's malfunction? Perhaps *B*'s *ignorance* of the chip's malfunction prevents *B* from defying *A*,

leaving *B* just as unfree as before. Modern prevention views do typically allow that ignorance can function as a freedom-limiting constraint in certain cases. As I now explain, however, prevention views can't coherently treat *B*'s ignorance as a freedom-limiting constraint in *this* type of case (and typically don't try to do so, for good reason).

Consider Kramer's characteristically sophisticated treatment of the relationship between ignorance and freedom. For him, not knowing *how* to do something can render one unfree to do it: if one doesn't know how to speak Italian,[17] for instance, or how to open a particular combination lock,[18] then one is unfree to do those things. Moreover, for Kramer the *extent* to which such ignorance reduces one's freedom depends upon the difficulty of overcoming it ('ignorance… amounts to a freedom-curtailing condition in proportion to the difficulty of its being overcome'), where difficulty is understood in terms of probability of success ('the probabilistic qualification attached to any ascription of freedom will reflect the likelihood of the dissolution of a person's ignorance if the person were indeed to endeavour (without any blunders) to achieve such a dissolution').[19] To take Kramer's example, this means that if the correct code for a combination lock were hidden somewhere within one's cell, one's freedom to open the lock would depend, in part, on the probability of one's finding it were one to attempt to do so.[20]

However, none of this helps the prevention theorist get to the conclusion that B suffers reduced freedom in the second period of *Heart Chip*. *B* doesn't lack any relevant instrumental knowledge in that period: the problem isn't that *B* doesn't know *how* to defy *A*'s threats, but that *B* doesn't know that it's *possible* to defy *A*'s threats (without untoward consequences).

---

[17] *The Quality of Freedom*, 265.

[18] 83.

[19] 266.

[20] 81.

Moreover, this ignorance is easy for *B* to overcome, in the relevant sense of 'difficulty', since *B* can test the heart chip at any time by defying *A* and seeing what happens. Were *B* to endeavour to overcome their ignorance in this way there is a very high probability that both *A* and *B* would shortly discover that the chip had malfunctioned, and that *B* was free to defy *A* without consequence.[21] So *B*'s ignorance does not, on Kramer's analysis, render *B* unfree.

Indeed, *B*'s situation is relevantly similar to that of Kramer's *Kathy*:

> Suppose that Bob has abducted Kathy and has placed her in a shed. Inadvertently, he leaves the door to the shed unlocked. A simple turn of the knob by Kathy would open the door and would thus enable her to escape altogether with very little effort. Nonetheless, because Kathy wrongly believes that the door is firmly locked…she does not even attempt to open the door, and consequently does not discover that she can readily escape. Her ignorance in this case does not deprive her of her freedom-to-leave-the-shed; instead, it simply leads her to fail to exercise that freedom.[22]

In the same way, *B*'s ignorance of the heart chip's malfunction doesn't deprive *B* of the conjunctive freedom of defying *A* and going on to act in ways that would be prevented by the activation of the chip; instead, it simply leads *B* to fail to exercise that freedom.

---

[21] Is *B* not prevented, by the risk of death, from *trying to discover* whether the chip is still functioning? No: the objective probability of death from the chip in the second period is zero, so there is no risk of death. Given this, is *B* not prevented by the *false belief* that there is a substantial risk of death from trying to discover whether the chip is still functioning? No: this false belief is, again, one that could be overcome simply by trying to discover whether the chip is still functioning and seeing what happens. That is, the same considerations that apply to the question of whether *B* is prevented from defying *A* also apply, recursively, to the question of whether *B* is prevented from overcoming their ignorance as regards their preventedness with respect to defying *A*.

[22] 265–66.

What's more, this is what any consistent prevention-based view of freedom *must* say about this kind of case. A unifying commitment of such views is their commitment to *preference neutrality*: the idea that facts about what a person is free to do are independent of facts about what that person wants to do. A supremely virtuous person, for instance, may have no desire whatsoever to act badly; indeed, the probability of their acting badly may be close to nil. On a nonprevention view, however, they nevertheless remain *free* to act badly inasmuch as nothing prevents them from doing so. Motivations and preferences are not themselves constraints. In the same way, if the shed door is unlocked, then Kathy is free to leave, and she is free to leave regardless of whether or not she is motivated to leave. Although her belief that leaving is impossible defeats any motivation she may have had to try to leave, her motivations are irrelevant to her freedom. That is, her false belief prevents her from leaving only insofar as it defeats her motivation to leave, and lack of motivation to leave is no constraint on leaving. Tempting as it may be to think that a belief in the impossibility of your doing *x* prevents you from doing *x*, therefore, such a principle cannot, on reflection, be accepted.

In sum: nothing prevents *B* from defying *A* (and living to tell the tale) in the second five-year period of *Heart Chip*. That's an important kind of freedom for *B* to have, even if *B* is unaware of it. Yet *B* nevertheless continues to act under coercion, and hence unfreely. In this further sense, *B*'s freedom remains diminished. To account for this, we must recognise that freedom requires the absence of coercion as well as that of prevention.

Perhaps the prevention-only theorist will respond by insisting that, since *B* isn't prevented from doing anything in the second period of *Heart Chip*, *B* isn't coerced into doing anything either, on the grounds that coercion requires prevention. Such an insistence would, however, be little more than a Procrustean attempt to reshape the idea of coercion in line with the prevention theorist's prior commitments. The fundamental *normative* point about *Heart Chip* is that *B* is every bit as subject to *A*'s will in the second period as in the first period, and

that there is therefore some significant sense in which *B* remains unfree in the second period. If the prevention theorist wishes to reserve talk of 'coercion' for cases involving prevention, then we will simply need another term to pick out the distinctive bad that *B* suffers continuously throughout both periods, that of subordination to *A*'s foreign will. This will not affect the main claims of this section, which are that there are two distinct kinds of disvalue at work in the case—blocked opportunities and interpersonal subjection—and that both are relevant to judgements regarding *B*'s degree of socio-political freedom. Although I maintain that it's perfectly proper to call the second of these 'coercion', this is not fundamentally a dispute about words.

2.2. Preference Manipulation

Not all forms of coercive interference involve threats. It's also possible to imagine cases in which behaviour is controlled by other illegitimate means. Consider:

> *Brain Chip: A* secretly implants a chip in *B*'s brain that allows *A* to alter *B*'s motivational states. For example, if *B* faces an exclusive choice between *x* and *y* and favours *x*, *A* can get *B* to choose *y* instead by using the chip to induce a stronger motivation in favour of *y* (or to weaken the motivation in favour of *x*). *A* routinely uses the chip to get *B* to comply with *A*'s wishes.

When *B* acts under *A*'s control in this way, *B* doesn't act freely, and *B*'s freedom is diminished.

Despite the absence of any external form of physical prevention, most contemporary negative liberty theorists agree that preference manipulation of this type can diminish freedom. For Kristján Kristjánsson, for instance, 'there is no morally relevant difference between preventing a man from entering a room by locking it up or by hypnotising him into staying

away from it. In both cases, the victim is unfree to enter, on a correct negative view.'[23] Benn, too, believes that preference manipulation, 'for example by media exploitation, advertising techniques, or religious and political ritual practices', can diminish freedom, since with the person's 'beliefs and preferences so managed, his actions can be controlled with no alteration to his objective choice conditions, effectively depriving him of the freedom to do what he nevertheless has the power or capacity to do.'[24] Similarly, Carter holds that 'forcible hypnosis and brainwashing can entail great reductions in freedom',[25] a position also endorsed by Kramer.[26]

Yet prevention theorists aren't entitled to this view. As I now argue, there are decisive reasons to adopt an analysis of cases like *Brain Chip* on which *A* doesn't remove any of *B*'s options, and therefore doesn't prevent *B* from doing anything. If I'm right, this means that prevention-only views of personal freedom can't account for *B*'s loss of freedom; and it suggests, again, that such views require supplementation by an independent account of freedom-as-noncoercion.

It's not unnatural to think that, in manipulating someone into doing (or intending to do) something, the mind-controller thereby *prevents* that person from doing (or intending to do) anything else, at least within some salient domain and at some level of description. As Kramer puts it: 'when the masters of [a person's] mind pre-decide her in favour of certain paths of conduct, they preclude her from opting for other such paths'.[27] Applied to *Brain Chip*, the idea

---

[23] *Social Freedom*, 100.

[24] *A Theory of Freedom* (Cambridge: Cambridge University Press, 1988), 154. Felix Oppenheim, similarly, holds that 'to indoctrinate someone is to limit his freedom' *Dimensions of Freedom: An Analysis* (New York: St. Martin's Press, 1961), 125.

[25] *A Measure of Freedom*, 206.

[26] *The Quality of Freedom*, 255–60.

[27] 257.

would be that in getting *B* to choose *y* by implanting a desire for *y* (and diminishing or eliminating *B*'s desire for *x*), *A* thereby prevents *B* from choosing *x*. The implanted desire for *y* thus functions as an obstacle, and constrains *B* just as surely as any external obstacle despite its internal location. However, while *some* forms of mind control may involve the creation of such internal obstacles, *preference manipulation* can't be understood in this way.

The problem is that one's own desires or motivations can't be treated as constraints on one's own possible actions.[28] Generally speaking, merely wanting to do something doesn't prevent one from doing something else. If I prefer apples to oranges, I'm not thereby prevented from choosing oranges; and if I later come to prefer oranges to apples, I'm not thereby prevented from choosing apples. Neither motivations nor changes to one's motivations are, in themselves, forms of prevention—if they were, then we'd always be prevented from doing anything other than that which we actually do. So when *A* implants in *B* a preference for *y*, it doesn't look as though *A* introduces a *constraint*. This isn't because it's internal to *B*, but because desires and preferences, in general, aren't the sorts of things that can constrain.

Perhaps some mind controllers implant non-standard desires. We might imagine, for instance, that *A* implants in *B* a preference for *y* that *causally necessitates B*'s choosing *y*. Assuming that preferences don't normally determine choices in such a mechanistic way, this

[28] This said, not all of one's desires and motivations may be 'one's own' in the relevant sense; see Harry G. Frankfurt, 'Freedom of Will and the Concept of the Person', *Journal of Philosophy* 68, no. 1 (1971): 5–20. In these cases, it may be possible to treat one's 'alien' desires as constraints on oneself narrowly construed, i.e. as constraints on one's 'true self'. Since the negative theorist is unlikely to be enthusiastic about incorporating a theory of the 'true self' into their analysis of socio-political freedom, however, I do not pursue this point here in any depth (though see my 'Agency and Inner Freedom', *Nous* 51, no. 1 (2017): 3–23). Nevertheless, it should be noted that not every illegitimately implanted desire must be alien in this sense, and we may therefore simply stipulate that, in *Brain Chip*, the alterations that *A* makes to *B*'s motivational states do not cause those states to violate the requirements for relevant ownership by *B*: see Gary Watson, 'Free Agency', *Journal of Philosophy* 72 (1975): 205–20; Harry G. Frankfurt, 'Three Concepts of Free Action', in *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), 47–57; Richard Double, *The Non-Reality of Free Will*, vol. 102, 3 (Oxford University Press, 1991), 31–61.

would constitute a special feature of the implanted preference. Moreover, it would be a relevantly special feature, for it would then be true of the implanted preference, but not of preferences in general, that it makes it physically impossible for *B* to choose *x*. In this case, therefore, it would be possible to hold that *A*'s interference prevents *B* from choosing (and therefore doing) *x* without having to hold, implausibly, that preferences are constraints in general.

However, manipulated preferences needn't causally necessitate choice in order to undermine freedom. As described, the chip generates in *B* a strong preference for *y* over *x* and this, presumably, raises the chances of *B*'s doing *y* considerably. Yet it needn't *deterministically* cause *B* to do *y*, just as an ordinary preference for *y* needn't deterministically cause *B* to do *y*. Hence suppose that, given *B*'s initial preferences, it's 99% probable that *B* will do *x*, but that the chip, once activated, makes it 99% probable that *B* will do *y*. If the chip is activated and indeed causes *B* to do *y*, then *B* is manipulated, and *B* acts unfreely. Freedom-undermining manipulation doesn't require deterministic causation.[29]

To flesh this out further: imagine, again, a person so supremely virtuous that the chances of their acting immorally in any given situation are at or close to nil. As previously noted, any nonprevention account of freedom must maintain that such a virtuous disposition is not itself a constraint on action, and that this person nevertheless remains *free* to act immorally

[29] This point is well established in the freewill literature. As Bernard Berofsky puts it: 'We are constrained to accept this conclusion on pain of conceding the possibility that no one has ever been manipulated, an absurd hypothesis. Although we can often feel supremely confident that a particular case of manipulation is also a case of determination, there is always at least a slight chance we are wrong. And should it turn out that a (prima facie) manipulated agent in situation *s* would respond under relevantly identical circumstances in the same way only 999 times out of 1000, would we really deny that he was manipulated in situation *s*?' ('Global Control and Freedom', *Philosophical Studies* 131, no. 2 (2006): 423). See also Alfred R. Mele, *Autonomous Agents: From Self-Control to Autonomy* (New York and Oxford: Oxford University Press, 1995), 139–43; Ishtiyaque Haji and Stefaan E. Cuypers, *Moral Responsibility, Authenticity, and Education* (New York: Routledge, 2008), 10–11.

absent other preventing conditions. Suppose, then, that *B*, who had previously been on course to choose *x*, is made (by activation of the brain chip) every bit as likely to avoid *x* as the supremely virtuous person is to avoid the immoral path. This is, clearly, an instance of freedom-undermining behavioural control. Moreover, its status as an instance of freedom-undermining behavioural control doesn't depend on any special assumptions about determinism or the nature of the implanted preferences. To the contrary, we may imagine that *A* inculcates in *B* an entirely normal motivational structure, one that is no more (nor less) effective in determining action than the moral orientation of the supremely virtuous person. By parity of reasoning, a nonprevention account of freedom should yield the result that *B* is nevertheless free to do *x*.

Indeed, this just follows from preference neutrality. The extent of one's opportunity set at a time doesn't depend on which of its members one is disposed to select, and nor does one become more or less free to do something merely by desiring to do it more or less strongly. Hence neither *B*'s original preference for *x* over *y*, nor *B*'s subsequent preference for *y* over *x*, should have any bearing on *B*'s freedom to do *x*. In illegitimately switching *B* from one preference to the other, therefore, *A* is meddling in states of *B* that are strictly irrelevant to judgements concerning *B*'s degree of preventedness.

If the implanted preference doesn't constitute a constraint in virtue of its *nature*, however, perhaps it does so in virtue of its *source*. Many nonprevention views of freedom distinguish between constraints that arise naturally and constraints that are imposed by other people, holding that only the latter compromise freedom.[30] This might be thought to offer an explanation as to why *B*'s implanted preferences function as impediments whereas *B*'s other, 'naturally arising' preferences don't. However, the distinction between natural and social constraints is of no help here, for two reasons. First, the point of the distinction is to restrict the

[30] See, e.g., Berlin, 'Two Concepts of Liberty'; Miller, 'Constraints on Freedom'; Kristjánsson, *Social Freedom*; Steiner, 'Individual Liberty'.

domain of freedom-relevant constraints, not to define the concept of constraint itself. If a light breeze doesn't constrain you, then it doesn't matter whether it's caused by natural atmospheric changes or by a person with a fan; it has no impact on your freedom either way. Similarly, if preferences aren't constraints, then it doesn't matter how they come about. The problem here is that preferences don't seem to be constraints in the first place, and the thesis that only socially-imposed constraints restrict freedom doesn't help with this.

Second, there are many ways to bring about motivational changes in others, and not all of them compromise freedom. Suppose, for instance, that *A* gets *B* to desire *y* by providing *B* with new information about the relative merits of *x* and *y*, and by pointing out good reasons for preferring *y*. This doesn't seem to involve any detriment to *B*'s freedom. So even if it were possible to show that externally inculcated preferences somehow constrain merely in virtue of their social aetiology, this would then prove too much, for it would also yield the implausible conclusion that rationally persuading someone to do something prevents them from doing otherwise.

Instead, the relevant difference between rational persuasion, on the one hand, and illegitimate modes of preference manipulation like that in *Brain Chip*, on the other, isn't that only one is a form of psychological *prevention*, but that only one is a form of *subjection to a foreign will*.[31] While nonprevention views are sensitive to the effects of interference on the opportunity sets of those interfered with, they're insensitive to the relation between wills constituted by such interference. To understand the loss of personal freedom in cases like *Brain Chip*, therefore, we need a way of thinking about coercion as a form of interpersonal subjection where this isn't just a special case of prevention.

[31] Elucidating this latter notion is obviously tricky. For attempts to do so in this context, see Gideon Yaffe, 'Indoctrination, Coercion and Freedom of Will', *Philosophy and Phenomenological Research* 67, no. 2 (2003): 335–356; Michael Garnett, 'Freedom and Indoctrination', *Proceedings of the Aristotelian Society* CXV, no. 2 (2014): 93–108.

3. The Ideal of Noncoercion

I've been arguing that there are two overlapping concepts of negative freedom, each valuable in its own right. One is the idea of non-prevention. Descended ultimately from Hobbes, and developed in great detail by contemporary negative theorists, this is, fundamentally, a matter of unconstrained opportunities. The other is the idea of non-coercion. Although also a prominent theme in the negative tradition, it's rare for it to be explicitly distinguished from the idea of non-prevention. This is, fundamentally, a matter of non-subjection to foreign wills. I've argued that to be a free person in the negative sense one must *both* be unprevented from performing a wide range of actions *and* uncoerced insofar as one performs any of those actions.

In making this argument I've relied on a somewhat general understanding of coercion. Providing a more specific account of the relevant conception is a large undertaking—at least as large as that of providing a comprehensive account of the relevant idea of prevention, the subject of several book-length treatments.[32] Fortunately, however, this paper's central claim—that nonprevention and noncoercion are distinct modes of negative freedom—can be made and defended without a fully comprehensive analysis of either. It requires only that we have a firm grasp on the normative *point* of the concepts of prevention and coercion, which is to pick out the distinctive bads of (respectively) blocked opportunities and interpersonal subjection.

Such talk of interpersonal subjection may seem to align the idea of coercion with that of *domination* as it figures in the republican tradition. Yet to equate the two would be to misunderstand both ideas. According to neo-republicans such as Quentin Skinner and Philip

[32] For a partial attempt to provide a more detailed account of coercion, see my 'Coercion: The Wrong and the Bad', *Ethics* 128, no. 3 (2018): 545–73.

Pettit, freedom requires the absence of domination.[33] This relationship is 'exemplified by the relationship of master to slave or master to servant':

> Such a relationship means, at the limit, that the dominating party can interfere on an arbitrary basis with the choices of the dominated: can interfere, in particular, on the basis of an interest or an opinion that need not be shared by the person affected. The dominating party can practise interference, then, at will and with impunity: they do not have to seek anyone's leave and they do not have to incur any scrutiny or penalty.[34]

As this passage makes clear, domination is not equivalent to interference: on Pettit's definition, domination is a *capacity* to interfere *on an arbitrary basis*.[35] Coercion, by contrast, is a mode of interference. Thus this paper's central point isn't the republican one that negative theorists ought to be concerned with unactualised powers of interference. It's a point about which types of interference ought to matter for negative theorists. It's therefore an 'in-house' dispute amongst negative theorists about how best to characterise the notion of freedom-relevant interference, one orthogonal to the central dispute between negative theorists and republicans. The claim is that both prevention and coercion ought to matter, and that they are both conceptually and evaluatively distinct: one, the evil of prevention, a matter of having obstacles put in one's way; the other, the evil of coercion, a matter of being forced to comply with another's will.

Hayek puts this point as follows:

---

[33] Quentin Skinner, 'The Third Concept of Liberty', *Proceedings of the British Academy* 117 (2002): 237–68; Philip Pettit, *Republicanism: A Theory of Freedom and Government* (Oxford: Clarendon Press, 1997).

[34] *Republicanism*, 22.

[35] 52.

> The question of how many courses of action are open to a person is, of course, very important. But it is a different question from that of how far in acting he can follow his own plans and intentions, to what extent the pattern of his conduct is of his own design, directed toward ends for which he has been persistently striving rather than toward necessities created by others in order to make him do what they want.[36]

Although coercion normally coincides with prevention, the two can sometimes come apart. An analysis of freedom that focuses only on the good of non-prevention will be incomplete insofar as it fails to capture the good of non-coercion, since a negatively free person is not only one who has unconstrained options, but also one who doesn't act under another's yoke.

[36] *The Constitution of Liberty*, 60–61.