



Norwegian University of Life Sciences
School of Economics and Business



Powers, Non-consent and Freedom

Stephen Mumford and Rani Lill Anjum

Working Papers No. 1/ 2014

ISSN: 2464-1561

Powers, Non-consent and Freedom

Stephen Mumford
University of Nottingham, Nottingham, UK
and Norwegian University of Life Sciences (NMBU), Aas, Norway

and

Rani Lill Anjum
Norwegian University of Life Sciences (NMBU), Aas, Norway

Abstract

There are a number of dispositionalist solutions to the free will problem based on freedom consisting in the agent's exercise of a power. But if a subject a is free when they exercise their power P , there is an objection to be overcome from the possibility of power implantation. A brainwasher, rather than directly manipulating a subject's movements, can instead implant in them a desire, to be understood as a disposition to act, and allow the subject to exercise such a power. It seems that, according to the dispositionalist theory of freedom, such an agent would still count as acting freely. There is a strong non-consent intuition that a is not free in such a case because they did not consent to having the power P – the desire in question. Filling out this intuition is not straightforward. But it can be done in terms of the exercise of P being regulated by higher-order powers of self-reflection. Such regulation is what allows an agent to either take ownership of a power or to reject it.

1. Free will and empowerment

It is very tempting to think that an agent acts freely when they exercise one of their causal powers. Such an intuition, we agree, is basically right. But it is not an entirely simple or clear-cut matter. There is a class of cases that challenges the view. These are where it looks as if the agent is exercising a power but the power that is exercised is not really regarded as the agent's own. And in these cases, the agent's belief, where there is one, that they are acting of their own free will is false.

There are a number of accounts now on offer in which the basis of free will is given in terms of the causal powers of agents (Kane 1996, O'Connor 2000, Vihvelin 2004, 2013, Ellis 2012, Lowe 2012, Steward 2012, Groff 2013b, Mumford and Anjum 2014, forthcoming). Let us call this the Powers Theory of Free Will (PTFW). Simplifying over the nuances of the various accounts, the core of this theory is something like the following:

(PFW): For any agent a and any power P , a acts freely when a exercises a power P that is under a 's control.

PFW is not being attributed to any particular author. It need not be. It is taken to be indicative of a common core of a number of recent contributions to the free will debate, all of which acknowledge the dispositions or causal powers of agents as one of the key notions in a solution to the problem. Such solutions would usually be classed as libertarian accounts of free will though some are not. The point, however, is that theories of free will such as PFW are all vulnerable to the type of objection we will present and develop here. We do so in order to show how PFW should respond. What we offer should thus be understood as a defense of PFW: a defense that centers on the clarification of what it is for a power to be the agent's own rather than anyone else's.

Let us first articulate some of the ideas behind PFW that will be useful later. The common core of all such accounts is the notion that agents are empowered, where this involves commitment to a metaphysics of causal powers (see Mellor 1974, Bhaskar 1975, Harré and Madden 1975, Ellis 2001, Molnar 2003, Mumford 2004). There are varying theories of what powers consist in. What is important for what follows is the understanding of many mental phenomena as powers or dispositions to act. Hence, in a tradition going back to Ryle (1949) but including Armstrong (1968) and more modern proponents such as Dennett (1991) and Mumford (1998), to have a belief, a desire, an emotion, or to have made a decision, all count as powers of an agent. We can call this a causal-dispositional theory of mind. A decision disposes someone to act in a certain way, for instance. An agent who has made a decision is thus suitably empowered in the sense that concerns us. Desires also dispose an agent to act though they may not lead directly to action if they are not accompanied by appropriate beliefs and other suitable desires. Conducive circumstances and indeed physical powers are also needed in order for an agent to act. Someone might desire to jump to the moon but know that it is physically impossible for them to do so, which rationally ought to be adequate grounds not to try to act upon the desire.

PFW says that, for free action, the power that is exercised must be one that is under the agent's control. One is not acting freely when one casts a shadow, usually; nor when one gravitationally attracts other massive objects. Such powers may well be powers of a , but they do not count as examples of a 's free action when a exercises them. The agent a has no control of such powers: a cannot choose whether or not to exercise them, for instance. Hence, if someone accidentally falls off a cliff, they could land on someone else, kill them, and survive the fall. The power to crush another person indeed belonged to the faller, and the power was exercised, but it is not thought of as a free action because the faller had no control over it. This point is usually taken as the basis of a moral and legal principle of responsibility. If a power cannot be controlled then it cannot be taken as a basis of praise and blame. For powers to be freely exercised, the agent must be able to choose whether and when to exercise them, at what time, to what degree, for how long, and so on (see for instance Aristotle's *Ethics*).

The notion of control may be thought to be doing too much work here, basically itself amounting to the attribution of free will to the agent. But we have to trust that it can be spelled out what control of a power amounts to without invoking the notion of free will itself. Our conclusion at the end of

the paper is a contribution to that very challenge. The choosing, in choosing when to exercise a power, is itself a power – a higher-order one – that the agent exercises in order to be free. It does indeed seem that we free agents have this power to choose, and we can often control when to exercise that power: when to make a choice. We had better ensure, therefore, that our account of free will as the exercise of a power is not circular.

We will assume that something like PFW has credibility and can be spelled out in an acceptable way. We can then turn to the matter at hand. There is a problem afflicting it that proponents of PFW have not adequately addressed. This is the problem that, for want of a better term, we will call the possibility of power implantation, which may be most easily comprehended in the case of implantation of a desire.

2. Power implantation

Suppose there were unscrupulous people who knew that PFW had been accepted as the correct account of free will and saw a loophole that they could exploit. It would create the possibility of the following kind of case.

An advertiser might realize that they are able to implant a desire for chocolate into a subject, which the subject then acts upon. It is often said that you can lead a horse to water but you can't make it drink. Instead of trying to force the horse, an easier solution might be to give the horse a desire to drink, if that could be done. Similarly, in advertising one doesn't oblige anyone to buy a product but one does try to place in them a desire for the product. The idea behind subliminal advertising is to hide the fact from the potential customer that they are being advertised to so that they then think of the desire as really their own when they exercise it. Should this be permitted? Can the advertiser say, invoking PFW, that when the subject quickly eats chocolate, having been exposed to the subliminal image, they are doing so of their own free will? They are, after all, exercising a power. On the powers account presumably beliefs and desires act as powers that the free agent chooses to exercise or not, as in classic causal theories of mind. Thus, a desire to eat chocolate, added to a belief about the nearest place in which it is sold, can cause an action of making a purchase.

Note that in this case we had best say that the desire is implanted into a subject rather than into an agent. If we used the term agent, it might look as if we prejudge that the subject is indeed free when they exercise those powers: that is, they are an agent with respect to those powers. This is precisely what is at issue. Are those truly the subject's powers that are exercised? If they are, it looks like they exercise agency. But it is arguably not agency if the power being exercised is in truth someone else's: the advertiser's.

Now let's move to a more dramatic case. Suppose an evil politician implants in the electorate a desire to vote for him, which they then act upon, and he subsequently wins the election. Did they vote freely? Was it a fair and valid election? We would almost certainly say not. Yet the politician's defense could well be that the voters acted freely because they exercised their powers and free will consists in PFW. Presumably the discussion hinges on the issue of whether it was really the voters' own desires that were being acted upon. The politician says so. He knows the voters had this desire

to vote for him. He knows because he implanted that very desire. And on Election Day, he merely left them alone to act on that desire. Yet the fact that he intervened to give them the desire – to implant it within them – suggests to us that it was not really their desire that was being exercised.

A final example is the most extreme. As we sometimes see in Cold-War spy stories, suppose the enemy side is able to brainwash and re-program a sleeper agent. Without that sleeper knowing, they have a disposition hidden within them to assassinate their own President when they hear the code word ‘elephantiasis’. The sleeper has no idea that they possess this disposition. They have no recollection of their brainwashing, the end stage of which was to remove all memory of the process. The enemy then need not intervene to kill the President. They merely make sure that the sleeper hears the code word, and the sleeper performs the assassination. Do they do so freely? PFW seems to suggest so. They exercised a causal power. And let us assume that they did so with some degree of control. They calmly purchase a gun, travel to the vicinity of the President, evade the security measures, select the perfect moment at which to fire, and so on. Would the enemy be right to place all the blame on the individual sleeper and accept none themselves? We tend to think not.

We could go on to trace out all variety of cases of brainwashing, indoctrination, hypnotism and wireless remote control of others. But the kind of case that is in mind is sufficiently established. Rather than directly control the subject, it seems that an evil manipulator, if they have the capability, would find it less complicated to implant a desire into the subject and then let that subject operate their own limbs in accordance with that power. The evil manipulator who knows the theory of PFW might also be able to protest their innocence. They might say that nothing they have done violates α 's free will as defined by PFW.

3. The non-consent intuition

There is a reasonable sense in which we want to resist this conclusion. The subject is not exercising ‘their’ powers at all, we might feel. They are acting upon someone else’s power that has been forced upon them without permission. This might be called the non-consent intuition. α has the power to F , and indeed F s, but α was not acting freely in doing so because α did not consent to the acquisition of the power, be that a desire, a belief, a decision, and so on. In that case, we shouldn’t really say that it is α 's power that is being exercised; therefore the case is no exception to PFW.

But can we fill out this intuition and make it a credible answer to the problem of implanted powers? We will see that it is complicated doing so in order to salvage PFW, but arguably it can be done. We will be urging, therefore, that PFW stands despite the objection. Explicitly, we will argue that an account can be defended of what makes a causal power an agent’s own and that implanted powers do not count as the agent’s. But this requires a focus back on what are the correct powers that relate to free will and how they can be used by a free agent to challenge any implanted powers.

A relevant issue is that while the subject of power implantation gave no consent to acquire the power in question, free agents rarely do give consent to power acquisition. We might then want to apply a notion of assumed consent; but when would we ever be in a position to assume that consent had been given if it is not given explicitly?

4. Which powers?

To solve the problem created by the possibility of implanted powers, we need to return to the basic theory and answer two questions. Free will consists in the exercise of powers, the theory tells us, but to make this idea immune to the objection from implanted powers we need to specify which powers and we need to know whose powers.

We already saw that a subject exercising a power of gravitational attraction was not a matter involving free will. So which powers do involve free will? Which are the powers whose exercise would count as free? Freedom in general, we would say, is a matter of being able to get what one wants: being able to exercise one's powers to satisfy objectives. And this includes physical powers, such as where someone jumps over a fence to get to the other side. But what we are considering here is free will, which is not quite the same as freedom generally. One can have free will without freedom, as in the case of a prisoner. Whether one can have freedom without free will is less clear but not a matter we need pursue.

In the case of free will, the concern is with the exercise of psychological powers, not that these can always be entirely separated from physical powers. The mind-body distinction is not always sharp. But the idea is that to have free will one must have a power to deliberate and make choices between options. On examination, however, these psychological powers are very likely to turn out to be complex. To be able to deliberate and make choices, one almost certainly must be able to form preferences, strategies for obtaining those preferences, and intentions. And to hold these powers, an agent arguably must be capable of perceiving and imagining: imagining some future non-actual state that is desired. This might require a power to represent, if representation is necessary for imagination (see the discussion in Molnar 2003, ch. 3).

What this shows is that free will should not be understood as a single power, as Hume's account of the will comes close to suggesting (1739: Bk 2), but through analysis of mental capacities, is likely to resolve into a cluster of many powers, probably hierarchically arranged. Such hierarchical and dispositional theories of mind have been developed in the past (Armstrong 1968) and one possible outcome might be that to have a power of free will requires a host of inter-dependent powers: a cluster or package (see also Kirk's 1994 basic package account of consciousness). Without a power to perceive, for instance, one cannot acquire information about the options between which a decision is to be made. Indeed, it is contentious whether one could even acquire any concepts with which to think about options. But a merely perceptive being might still not be able to reason about those options and select one. It seems we have to allow, then, that a number of these powers will be pre-conditions of having free will. It might be that the having of free will requires a complete package.

We will not pretend that we can resolve all these issues here, nor indeed state definitively what the list of causal powers in that basic package for free will must be. It is clear that these questions can be answered finally only once we have a complete theory of mind. And we have also to allow that free will might be a complex power that can be variably realized by different packages of psychological powers, where each package would then be a sufficient but not a necessary condition of free will.

For ease of future reference in what follows, we propose to call this bundle of powers, which is the basis of free will according to PFW, the *causal will powers*. And we then want to say that an agent acts freely when they are able to exercise their causal will powers. We mean here the whole package, for it may be possible to exercise one of the powers within the cluster – the perceptive power, for instance – that on its own does not count as an act of free will. But, when the whole package is available, then in virtue of that a bearer has the power of free agency. There is, we will argue later, a particularly special power within this cluster of causal will powers. But we will not detail it until we come to the question of whose powers, where it will play an important role.

Much of the discussion of free will has centered on whether our decisions have causal antecedents. But this, we maintain, is not the issue as long as the agent has causal will powers. What counts is that an action attempts to realize the agent's desires, decisions or choices. It matters little, except in special cases discussed below, if those desires have causal antecedents. Suppose dehydration gives Peter an urge to drink. Peter can still choose freely to drink even though the heat and then his biology were in part responsible for the desire Peter is satisfying in drinking. Similarly, many of Jane's desires are provided by her socio-historical context. But they are still her desires nevertheless and the ones she aims to achieve in her choices and actions.

It is significant that the causal will powers are powers of agents specifically: not of their limbs or their molecules. Certain powers are applicable only above a particular level in nature and agency is an excellent example. Neurons don't act or even think. Eyes don't perceive. Persons think and act (and possibly higher-level collective entities can do so too). Thus, the freedom of the causal will powers is the agent's freedom, not that of their body or its neurons.

Our second question, then, is the one of whose powers are being exercised, which is so vital in addressing the problem of implanted powers. However, before approaching this, more features of a powers-based theory of free will need to be outlined as they will be essential for understanding the attraction of the position eventually advocated.

5. Supplementary theses

A powers-based theory of free will, to have plausibility, ought to come with a number of caveats. Supplementary theses should be added to PFW that make it considerably more plausible.

First, we should not commit to an overly-sophisticated theory of mind. We have to acknowledge *automaticity*. A number of free acts are automatized: for instance, most of the actions performed in driving a car, riding a bicycle, doing the dishes and so on. We should not, then, have a theory of free will that requires too high a standard for behavior to count as free; for instance, which states that only those acts are free that have been performed after distinct conscious episodes of deliberation and decision. This could have the upshot that an agent exercises their free will only a few times a day. When one turns on the indicator while driving, however, one might not even think about it, let alone form a desire, a plan for its realization and then an intention to act. Are automatized actions any less free? One might think that a strong swimmer, who can do so with little attention to the details of their stroke, is just as free, if not more so, than one who has to think about every motion.

Mastery of an action seems often to be a matter of automating it and so it would be a perverse theory of free will if one thereby lost one's freedom and agency with respect to it.

But there is an aspect of any dispositional theory of mind that the proponent of PFW should really seek to exploit. To be in possession of a power is to be able to do something. An agent is able to deliberate about any action they freely perform but they need not actually do so. Not actively deliberating should not in itself prevent an action from counting as free as long as one is able to do so, if needed. Hence, our driver, if asked whether they are really sure that they want to deploy the indicator, is certainly able to think carefully about it and deliberate. Similarly, when washing the dishes one might concentrate on one's actions when it comes to a particular valuable piece of china. But in the vast majority of cases we tend not to concentrate that much. Automaticity is fine for us; indeed a powers-based theory of free will sits especially well with it.

Second, PFW ought to invoke what is now called the dispositional modality (Mumford and Anjum 2011: ch. 8). If the world is one of necessity, it is regarded as a threat to our freedom. If prior states fix or determine the future, it seems that we have no genuine choice and free will is an illusion. This is why free will is regarded by many as incompatible with determinism (van Inwagen 1983, ch. 3, Mumford and Anjum 2014). But suppose the world is one of pure contingency or chance. Then we seem to be in no better a position. If, given prior states of the world, anything at all could follow, then we seem to have lost control (see Watson 1982: 2). In an entirely indeterministic world, one seems to be a slave to chance: and this is no better than being a slave to necessity. Merely showing the falsity of determinism is not enough to establish the existence of free will, therefore. Indeterminism can be just as big a threat and the challenge seems to be to show how free will is compatible with indeterminism.

The idea of the dispositional modality, however, is that the modality of causation is neither that of Hume's pure contingency nor that of necessity. The latter is crucial. Hume's opponents have often defended necessary connections between cause and effect, perhaps because that is what Hume attacked (see, for example, Harré and Madden 1975 and Ellis 2001). But then this suggests that, for free will, agents would have to step outside the web of natural necessity. That agents are able to do so would be a brave position to defend, though some forms of libertarianism seem aimed at this kind of solution. Instead, a cause should be taken as something that tends towards an effect, and often succeeds in producing it. Because the dispositional notion of causal production is not the same as causal necessitation, we can say that being causal patients – having causes act upon us – does not thereby necessitate our actions. But nor does that leave us with pure contingency, which would rob us of all control over our actions. One's actions still tend to produce their desired effects, even though they cannot guarantee them, and this is enough for us to ascribe responsibility for outcomes to that agent (Mumford and Anjum forthcoming).

A third supplementary claim to PFW follows from this. It is vital that PFW commits to a realist theory of causation in order that the agent has free will. They must cause their own behavior through the exercise of their real causal powers. Not every dispositional account of free will has been so explicit on this point, partly because there is a reductive conditional analysis of dispositions that still attracts some (see for instance Berofsky 2011). To offer a conditional analysis of dispositions is to deflate them to the point where they can no longer do the required work for the theory. As Groff (2011a,

ch. 6) argues, genuine agency requires realism about our causal powers. If our actions were only contingently conjoined with outcomes, or our desires only contingently conjoined with our actions, we would have no good reason to be interested in them and no responsibility. Humean accounts of causation and powers suggest that constant conjunction, with temporal priority and contiguity, is all the causation we need. But, if that were all the causation we had, we would have no rational basis for action. Our decisions would have no connection with our behavior and our behavior would have no connection with any outcomes. Those outcomes could not truly be thought of as consequences of our actions. The problem of induction is particularly acute, given Hume's account, and strikes agency as much as any other causal notion. There would be no rational basis for inductive inferences about our own behavior.

A fourth and final supplementary point is that a dispositional theory of free will ought to make the most of the scalarity of dispositions. They come in degrees. A wineglass and car windscreen may both be fragile but the former to a greater degree than the latter. A dispositional theory of causation, for example, makes play of this point (Mumford and Anjum 2011). And this has clear implications for the free will debate. Many discussions of free will talk as if it is an all or nothing matter. There may well be many things that have no free will at all, panpsychism aside. But of those things that have free will, some have more of it than others. A dog, for instance, could have a degree of free will but perhaps not as much as a human being. Humans probably have more of the powers that are the basis for free will than dogs have, we suspect. And if rational faculties were a precondition of free will, then we ought to allow that some people have more free will than others, given that some are more rational than others. It also allows the possibility to say that there are some things that impinge on one's freedom – perhaps advertising or peer pressure counteract one's liberty – but without preventing it altogether. The scalarity of dispositions should be found in the scalarity of free will, therefore, and the importance of this point cannot be emphasised enough in PFW.

6. Whose powers? Initial solutions

We now address the issue of proper ownership of a power, especially in the light of the power implantation cases. What they show is that merely bearing a power is not enough to make it an agent's own, for it may be a power that was implanted without consent.

The initial thesis can be articulated as:

T⁵: We are free (i.e. have free will) when we are able to exercise our causal will powers.

This raised the question of what makes the exercised powers *ours*? What if a brainwasher gives a subject a desire, for example to assassinate the President? Was the subject free in doing so? Arguably not: because it was not really their desire that was being acted upon. But how can we pin

down exactly what it would mean for it to be a subject's own power that is being exercised? We will now examine some relevant options before settling on our answer, which will be thesis **T**.

Webber (unpublished) has a view of what makes something the agent's power rather than the brainwasher's. One's true dispositions are those that are integrated with the rest of one's character. Hence, in the sleeper agent case, it is quite out of character for this person to suddenly go out and perform an assassination. They have never killed anyone before, we assume, and if they are ever questioned on matters of violence, they avow pacifism. The assassination-desire stands out as one that is at odds with all the other dispositions of this person and this is what shows that it is not really one of their own but has been artificially implanted in them. We may articulate the view as follows:

T⁴: We are free when we are able to exercise the causal will powers that are integrated with the rest of our character.

However, **T⁴** is inadequate. There are two reasons for this. First, it would only be able to capture those implanted powers that did indeed clash with other aspects of a subject's character. It is possible that someone implant a power that a person was previously lacking but which is nevertheless perfectly consistent with all a person's other powers. For example, suppose all the various desires and beliefs of a person are equally consistent with a liking for both the color turquoise and the color orange. This person's character no more favors a liking for one than the other. A fashion designer might implant in this person a preference for turquoise, which then appears integrated in that person's character. The subject begins to buy turquoise clothing. Arguably, it is not their own powers alone that are being exercised in making these purchases.

Second, **T⁴** seems to rule out the possibility that one might choose freely to act out of character. Such a possibility is vital to a libertarian account of free will because it might be thought that acting out of character is the ultimate expression of free will. Perhaps such thinking is behind Raskolnikov's decision to commit murder in Dostoevsky's *Crime and Punishment*. One may think of there being some acts the performance of which is motivated simply because it would be an assertion of one's freedom. Any fixity of a person's character could be seen as a constraint on this. Less dramatically than murder, someone might see there is something in their character that they wish were not there and try to change it. Perhaps someone doesn't like that they are shy. They might then agree to give a public talk, completely out of character, and we ought really to grant that this can be a free action on their part. Indeed, one really should allow that one is most free if one is able to overcome constraints of character. **T⁴**, therefore, seems to rule some acts as free when they don't seem to be so (the preference for turquoise) and some as not free when it appears that they are (acting out of character).

We need another answer. Here is a suggestion: one's real dispositions are those that have been acquired in the right way, for example, learning to play the violin through a series of lessons. A brainwasher planting a disposition directly in someone's mind does not count as the right way of acquiring a power. There is something jarring, sudden and discontinuous about this mode of

acquisition, whereas violin lessons lead to a gradual acquisition of the ability. We can articulate this as follows:

T³: We are free when we are able to exercise the causal will powers that have been acquired in the right way.

The problem here is what counts as a power having been acquired in the right way. Both the brainwasher and the advertiser implant in someone a power or desire. It looks as if these are cases of forced acquisition of powers. But there are some dispositions that we are forced to have yet still feel free when we exercise them. A child can be thrown in a swimming pool reluctantly at age three, for instance, and learn to swim from doing so. If asked, the child may state explicitly that they are not interested in learning to swim. Yet, if forced to learn, they can nevertheless be acting freely on the subsequent occasions that they exercise the power, later in life. Forced acquisition is not, therefore, enough on its own to distinguish powers that do not truly belong to a person even if they have them and exercise them.

The real difference here, it might be thought, is that the power to swim is an enabling power, in one's rational self-interest, whereas the power to assassinate someone is not: neither is the desire to drink more alcohol or eat more chocolate. And this means, one might think, that a rational agent should have wanted the power to swim and not the others. Thus, if one asks the grown woman whether it was right that she was forced to acquire the power to swim when a three-year old, against her wishes, she should say yes. She might indeed go on to force her own children to learn to swim, knowing that it is good for them to be so enabled. The ability to swim is in one's rational self-interest because it is empowering, whereas the powers to kill or eat junk food are irrational because ultimately they are disempowering. This line of thought might tempt us towards the following view:

T²: We are free when we are able to exercise the causal will powers that are in our rational self-interest.

Again, though, it is easy to think of a problem with **T²**. It automatically rules out that someone could freely act irrationally: for instance, that someone may smoke or drink and do so by free choice. There is of course a tradition in which one is thought to be free only insofar as one follows reason. Kant's categorical imperative might be offered as an example. But we are not working within that Platonic-Kantian tradition here. We want to allow that someone who does wrong or acts irrationally might still be doing so of their own free will. This applies to actions against the agent's own rational self-interest. We are effectively divorcing the issue of freedom from any matters of ethics, including the question of self-interestedness. Without that, anyone who has done wrong could automatically claim as a defense that they did not do so freely.

The smoker might then say that they realize smoking is bad for them but that it is their own power that is being exercised when they smoke because, although smoking is bad, it is nevertheless a causal power that they want to have, and indeed want to exercise. This leads us on to thesis T^1 :

T^1 We are free when we are able to exercise the causal will powers that we want to have.

We are almost at an acceptable theory of what it is to properly have a causal will power; but not quite. Even this will not do. Again, it is easy to think of counterexamples. The addict may have a desire for heroin and, even when questioned, say that they have no desire to give up this desire. Indeed, T^1 does not even help in the original power-implantation cases. Once the sleeper agent's power to assassinate has been activated, they may profess that they are perfectly happy in exercising and having the power to kill. Similarly, following subliminal advertising, a subject may feel perfectly content with their desire to eat chocolate. That, of course, is the whole problem of power implantation. The subject thinks of the desire as their own and is perfectly happy to accept ownership of it, even though we have information that suggests it is not really theirs.

We cannot accept any of T^5 - T^1 , therefore. They all have counterexamples. And yet they all told us something that seemed to be important for power-ownership. What we really need, then, is an account that includes what was plausible about each of T^5 - T^1 but does not contradict important intuitions about free will.

7. Thinking about your powers

In answering the question of whose powers, we must return to the issue of which powers. This is because it is vital to the proper ownership of an agent's powers that they are able to scrutinize them. And this is to have a higher-order power – a self-reflective one – to think about one's first-order powers and duly moderate or regulate them. Hence:

T : We are free when we are able to exercise our causal will powers that can be regulated by higher-order powers of self-reflection.

Frankfurt (1971) has already invoked such higher-order powers (i.e. second-order desires). Steward (2012), on the other hand, downplays them. For the issue of proper ownership, however, they take on an important role, especially when accompanied with the correct dispositional account of agency, as described above. We should, therefore, motivate the requirement of self-reflective powers.

A wanton is someone who acts on their desires without self-reflection: without thinking about the morality or rational self-interest of those desires. They may desire a chocolate bar, for instance, and

immediately seek to have the desire satisfied, without any consideration of the negative effects. Similarly, a psychopath may be unable to consider the effects of their actions on others and thus move straight from desire to action. Now, bearing in mind the important point that freedom can come in degrees, one might not want automatically to say that a wanton has no freedom at all but they might not have freedom in the highest degree because they lack something that is, if we are right, able to give us so much more freedom.

With self-reflection, a number of new possibilities are open for an agent. One could choose freely to act out of character, for instance. A free agent is aware what their desires are and is able to evaluate them. To conquer shyness, perhaps, or in an extreme adverse situation, he or she may choose to act out of character, seeing that their natural inclinations should be resisted. Similarly, an adult human, capable of self-reflection, is correctly held responsible for their smoking, even though it is against their rational self-interest. They are considered capable of reflecting on their desire and understanding the negative consequences. If they decide to smoke nevertheless, then we hold them responsible. Among other things, it is reflection that allows us to see the difference between enabling powers and those that are disabling. Again, we should allow this self-reflective power to be one that can come in degrees. Children may acquire it gradually. A five-year-old could strike an older sibling and then be admonished. The defense of the wanton could be employed: they didn't realize it was wrong. But a reply might be that they should have realized. They now have enough rationality to reflect on their desire to strike out, and see that it is wrong, and we are therefore holding them responsible.

There is a possible objection: the sort of counter-example we raised against T^5 - T^1 . The brainwasher could implant even the higher-order, self-reflective powers, which are not then the subject's own. But this does not really undermine T . Those higher-order reflective powers are ones that any rational agent would want since they are constitutive of freedom. Hence, the brainwasher could only be giving his victim something that enables her freedom and something that she should have already. In that case, the brainwasher would merely be overdetermining the subject's freedom. For example, if the brainwasher implants in a subject a desire to kill children, they undermine that desire if they also give that subject a higher-order power to reflect morally, rationally and self-consciously on that first-order desire. The self-reflective power has every chance of spotting this rogue desire to kill and allowing the subject to disown it. What is more, the power of self-reflection is one that when itself reflected upon can be understood as enabling freedom. Thus, even if an agent discovered that their self-reflective power has been implanted in them rather than acquired in the natural way – through normal human development – it is a power that they would be right to accept ownership of. It would be like the acceptance of a gift (of freedom).

Having just mentioned the idea of reflecting upon the ability to self-reflect, the question arises of whether there are third-order powers and powers of even higher-orders. It would perhaps be useful if there were. One could then reflect on any relatively lower-order power and decide whether or not to take ownership of it. But might this require too sophisticated an account, possibly an infinitely ascending one? Does one need an infinite number of ever higher-order powers in order to have one free lower-order power? Arguably not. There might just be one single power that is needed: what Hofstadter (1979: ch. 15) calls an ability to 'jump outside the system'. Thus, for any set of thoughts we entertain, we are able to pause, bracket those thoughts, and have thoughts about them. We can

do this even if the set of thoughts itself includes higher-order thoughts. Any collection of thoughts can be thought about. But it is the same ability involved in all cases – that is exercised upon a smaller or larger set of thoughts – rather than a wholly new ability of an even higher-order.

T therefore stands but it still requires some explanation. There is, for instance, no assumption regarding the moral theory against which self-reflection occurs, so we are not automatically moving back to the Platonic-Kantian tradition where only the good is free. A racist might act reflectively on their racist values and adjust their behavior in accordance with them, for example. Finding a first-order inclination to behave decently towards someone of a different race, they could reflect on it with respect to their background values and counter their initial inclination. People can be both free and immoral.

And we should also note the dispositional character of **T**. It does not say that every decision and every first-order power that is exercised is indeed reflected upon, only that the free agent is capable of doing so. It is a power at their disposal – an ability – and part of their practical rationality consists in good judgment when a first-order power should be reflected upon. In some circumstances, it might be considered a moral failing if someone acts without reflection: seizing a large slice of cake, for instance, without even considering whether there is enough for everyone else. **T** is not being overly rational about what actually goes on within each free-person's thinking. But merely the ability to self-reflect, we contend, is sufficient for an account of power-ownership. It allows the possibility, for instance, of reflecting upon an implanted power and deciding to jettison it, as one should with the desire to assassinate. Or one can accept the power and thereby take ownership of it, as with the power to swim.

It is worthwhile noting, however, that we should also allow that there are cases of weakness of the will or akrasia. After reflection, the akratic person might still do what they have concluded to be wrong. But, as the term 'weakness of the will' suggests, akrasia detracts from one's freedom: from one's causal will powers. The akratic loses some of their freedom to act as they think they should; and is thus losing control. A strong-willed person, in contrast, is one who is more able to do what they think is right, appropriate or required.

8. Conclusion

Power implantation is a rather fanciful possibility. Perhaps it could happen. Regardless of that, our consideration of the issues has brought to light something important about what proper ownership of a power requires if it is to be a power regarded as freely possessed and exercised. Lots of things have powers, such as machines, inanimate objects, substances, plants and animals. A realist about powers has a more active view of the world than has a Humean. The power of liquid to dissolve sugar really is in the liquid and to that extent the power is self-directed. But this only serves to show us that free will must consist in more than simply the exercise of a causal power, otherwise we would be prepared to ascribe it to cuckoo clocks.

We have argued that free agency requires a special kind of causal power, one the possession of which allows for the proper ownership of powers. It is a self-reflective power to think about, assess

and evaluate one's first-order powers. The higher-order power is a good basis for legal and moral responsibility, on the assumption that these require free will. Unless one is able to consider one's powers, one is unable to move beyond the position of the wanton and attain full freedom. This is the sort of account the defender of a powers-based theory of free will ought to hold.

Acknowledgments

Earlier versions of this paper were presented at the universities of Konstanz, Granada, Reading, Rennes, Sheffield and NMBU Aas. We are grateful to all who provided comments but in particular to Jonathan Webber for discussion, ideas and for his project with Alberto Masala on 'The Architecture of Personal Dispositions', at which related work was presented.

References

- Aristotle, *Ethics*, trans. J. Thomson, London: Penguin, 1953.
- Armstrong, D. (1968) *A Materialist Theory of the Mind*, rev. edn, London: Routledge, 1993.
- Berofsky, B. (2011) 'Compatibilism without Frankfurt: Dispositional Analysis of Free Will', in R. Kane (ed.) *The Oxford Handbook of Free Will*, 2nd ed., Oxford: Oxford University Press, pp. 153-74.
- Bhaskar, R. (1975) *A Realist Theory of Science*, Leeds: Leeds Books Limited.
- Dennett, D. (1991) *Consciousness Explained*, London: Penguin.
- Ellis, B. (2001) *Scientific Essentialism*, Cambridge: Cambridge University Press.
- Ellis, B. (2012) 'The Power of Agency', in R. Groff and J. Greco (eds) *Powers and Capacities in Philosophy: the New Aristotelianism*, New York: Routledge, pp. 186-206.
- Frankfurt, H. (1971) 'Freedom of the Will and the Concept of a Person', in *The Importance of What We Care About*, Cambridge: Cambridge University Press, 1998, pp. 11-25.
- Groff, R. (2013a) *Ontology Revisited*, London: Routledge.
- Groff, R. (2013b) 'Whose Powers? Which Agency?' in R. Groff and J. Greco (eds) *Powers and Capacities in Philosophy: the New Aristotelianism*, New York: Routledge, pp. 207-27.
- Harré, R. and Madden, E. H. (1975) *Causal Powers: A Theory of Natural Necessity*, Oxford: Blackwell.
- Hofstadter, D. (1979) *Gödel, Escher, Bach: an Eternal Golden Braid*, London: Penguin.
- Hume, D. (1739) *A Treatise of Human Nature*, L. A. Selby-Bigge (ed.), Oxford: Clarendon Press, 1888.
- Kane, R. (1996) *The Significance of Free Will*, New York: Oxford
- Kirk, R. (1994) *Raw Feeling*, Oxford: Oxford University Press.

- Lowe, E. J. (2012) 'The Will as a Rational Free Power' in R. Groff and J. Greco (eds) *Powers and Capacities in Philosophy: the New Aristotelianism*, New York: Routledge, pp. 172-85.
- Mellor, D. H. (1974) 'In Defense of Dispositions', *Philosophical Review*, 83: 157-81.
- Molnar, G. (2003) *Powers: A Study in Metaphysics*, S. Mumford (ed.), Oxford: Oxford University Press.
- Mumford, S. (1998) *Dispositions*, Oxford: Oxford University Press.
- Mumford, S. (2004) *Laws in Nature*, London: Routledge.
- Mumford, S. and Anjum, R. L. (2011) *Getting Causes from Powers*, Oxford: Oxford University Press.
- Mumford, S. and Anjum, R. L. (2014) 'A New Argument against Compatibilism', *Analysis*, 74: 20-25.
- Mumford, S. and Anjum, R. L. (forthcoming) 'Freedom and Control: the Modality of Free Will', *American Philosophical Quarterly*.
- O'Connor, T. (2000) *Persons and Causes*, New York: Oxford University Press.
- Ryle, G. (1949) *The Concept of Mind*, London: Hutchinson.
- Steward, H. (2012) *A Metaphysics for Freedom*, Oxford: Oxford University Press.
- Van Inwagen, P. (1983) *An Essay on Free Will*, Oxford: Oxford University Press.
- Vihvelin, K. (2004) 'Free Will Demystified: a Dispositional Account', in P. Russell and O. Deery (eds) *The Philosophy of Free Will*, Oxford: Oxford University Press, 2013, pp. 166-89.
- Vihvelin, K. (2013) *Causes, Laws, and Free Will*, Oxford: Oxford University Press.
- Watson, G. (1982) (ed.) *Free Will*, Oxford: Oxford University Press.
- Webber, J. (unpublished) 'Character and the Self'.