

# Freedom, Coercion and Discursive Control<sup>1</sup>

RICHARD HOLTON

If moral and political philosophy is to be of any use, it had better be concerned with real people. The focus need not be exclusively on people as they are; but it should surely not extend beyond how they would be under laws as they might be. It is one of the strengths of Philip Pettit's work that it is concerned with real people and the ways that they think: with the commonplace mind.

In this paper I examine Pettit's recent work on free will.<sup>2</sup> Much of my concern will be to see how his contentions fit with empirical findings about human psychology. Pettit is a compatibilist about free will: he holds that it is compatible with determinism. But he finds fault with existing compatibilist accounts, and then proposes his own amendment. My aim is to challenge his grounds for finding fault; and then to raise some questions about his own positive account.

## STANDARD COMPATIBILIST ACCOUNTS & PETTIT'S CRITICISMS

The standard compatibilist accounts of free will derive from Hobbes. From him and those who followed him we get a very simple picture: roughly, to be free is to act (or, perhaps, to be able to act) on one's desires. Elegant and enduring thought this account is, it succumbed to the realization that it lets in too much. Addicts, for instance, are paradigms of those who lack free will, and yet typically they act on their desires, desires for the objects of their addiction. The Hobbesian account needs to be restricted.

There are two main forms that the restrictions have taken in the subsequent discussion. Firstly, we might restrict the desires upon which the free agent acts. We might, as in Frankfurt's early writings, restrict them to the desires that the agent desires to have. Or we might, as in Frankfurt's later writings, restrict them to the desires with which the agent identifies. In Pettit's terminology such accounts posit *freedom as volition*.

Alternatively, we might think that desires are not enough. We might try to add a cognitive component, to give us what Pettit calls *freedom as rational control*. We might require that free agents perform those actions that they believe to be most valuable; or those that they believe to be best in some more open-ended way. Or, more demanding, we might require that free agents get (or be able to get) their beliefs *right*. We might require at least that they be *rationally formed*. In addition, we might require that they be

---

<sup>1</sup> Thanks to the audience at the Pettit conference for an excellent discussion; and to Geoff Brennan, Rachana Kamtekar, Rae Langton, Philip Pettit, Michael Ridge Tim Scanlon and the referees for OUP.

<sup>2</sup> *A Theory of Freedom* (Cambridge: Polity, 2001).

*true*: we might require that free agents respond (or be able to respond) to the reasons that they actually have, and not just to those that they believe themselves to have.

Pettit thinks that both these kinds of account are vulnerable to the same fundamental failing: their inability to deal with hostile coercion. Suppose that someone is threatening to do something very nasty to you unless you do as they say; and suppose that you quite rationally yield to their threat. Have you acted freely? It seems that both of these accounts will say that you have. You have acted on your strongest desire, a desire with which you fully identify; and your action is, by hypothesis, rational.<sup>3</sup> This is the conclusion that many compatibilists have drawn: coerced acts are free acts, at least in the sense of ‘free’ that the compatibilist is after. Yet there are two kinds of intuition that make this conclusion questionable:

(i) *Metaphysical intuitions*: coercion is exactly the kind of case in which we would normally deny that you acted freely, or willingly, or of your own free will.

(ii) *Moral intuitions*: coercion affects moral responsibility. If you do something as a result of the coercion that hurts some third party, then we will ordinarily think of you as less culpable than we would if you had caused that hurt without the coercion. Sometimes, though certainly not always, we might think of you as not culpable at all. Yet moral responsibility standardly requires freedom; so a plausible explanation of why you lack responsibility is that you lack freedom.

Pettit’s contention, based upon these sorts of consideration (though he does not distinguish them as I have) is that the standard compatibilist accounts are inadequate. Any plausible compatibilism must entail that coercion removes freedom. Of course, this is a thought with which many writers have agreed. The idea is a mainstay in discussions of political liberty.<sup>4</sup> Pettit’s aim is to bring together this political literature with that of the compatibilists, focussed as the latter are on personal responsibility. The claim is that underlying both discussions is a single notion of freedom, one that is intimately linked to moral responsibility. According to this single notion, one is free just in case one is fit to be held responsible. Pettit argues that by linking the two literatures in this way we increase the constraints on an acceptable account of freedom, and thus

---

<sup>3</sup> Actually it is far from clear that we simply assert that there are cases in which succumbing to coercion is rational; but I shall not pursue the issue.

<sup>4</sup> It is also there in some of the compatibilist literature: that which stresses the idea that free actions are those that are unconstrained. For an influential example, see Ayer, ‘Freedom and Necessity’ in *Philosophical Essays* (London: Macmillan, 1954) 271–84. Other compatibilists, whilst following a basically Hobbesian line, have tried to provide further conditions that free actions must meet, conditions that coerced actions will fail. Gideon Yaffe argues plausibly that Locke falls into this class (see *Liberty Worth the Name*, (Princeton, Princeton University Press, 2000), Chapter One); another example is Frankfurt, ‘Coercion and Moral Responsibility’ in *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988) pp. 26–46, which I discuss below.

eliminate accounts that seem otherwise plausible. He proposes an account that adds a condition of *discursive control*.

I shall postpone discussion of Pettit's positive account until the second part of this paper. In the first I shall focus on his criticisms of the traditional account. For I am sceptical. My suspicion is that the kind of freedom that is typically violated by coercion is very different to the kind of freedom that is violated in the cases that form the mainstay of the compatibilist literature: cases of thought manipulation, automatic action, post-hypnotic suggestion and the like. And I think that the way in which coercion lessens moral responsibility is very different from the way in which moral responsibility is lessened in those cases. Indeed, I think that the main basis for our metaphysical intuition that coercion compromises freedom is quite different from the main basis of our moral intuition that it reduces responsibility. It is exactly the tendency to think that our metaphysical and moral intuitions about coercion must have a single common source that has impeded understanding. In short, my thesis will be this:

(a) the reason why we think that coercion compromises freedom is because it removes our autonomy, where this is understood as freedom from manipulation by others. If our autonomy is violated we remain agents, but agents whose actions are being manipulated.

(b) the reasons why we think that coercion lessens moral responsibility are varied, but they typically involve treating the coercion as a justification or excuse; we accept that the agent had control of their action, but we think them justified in doing so, or excuse them if not. Whilst the fact that the agent was manipulated has some bearing on their justification or excuse, it is far from the main factor.

In contrast, the notion of free will that the compatibilist is trying to capture is the idea that stands behind agency itself. The person who is moved by post-hypnotic suggestion or brain interference ceases to be an agent at all.<sup>5</sup> The loss of moral responsibility follows directly from the loss of agenthood; there is no need for justification or excuse, since there is no action to be justified or excused. To make the case I need to look at coercion in much more detail. I start by examining the issues that surround our metaphysical intuitions.

---

<sup>5</sup> Or at least, post-hypnotic suggestion as philosophers conceive of it. The reality looks to be far more responsive to the existing attitudes of the subject. See, for instance, A. Barnier and K. McConkey, 'Posthypnotic Responding away from the Hypnotic Setting', *Psychological Science*, 9 (1998) 256-262.

## METAPHYSICAL INTUITIONS

It is a feature of coercion as we ordinarily understand it that it is the result of an action by another agent: you can be coerced by a person, but not by nature. Yet philosophers have been keen to minimize the philosophical importance of the point. Harry Frankfurt writes:

Only another person can *coerce* us, or interfere with our *social* or *political* freedom, but this is no more than a matter of useful terminology. When a person chooses to act in order to acquire a benefit or in order to escape an injury, the degree to which his choice is autonomous and the degree to which he acts freely do not depend on the origin of the conditions which lead him to choose and to act as he does. A man's will may not be his own even when he is not moved by the will of another.<sup>6</sup>

There is perhaps something right about this as a moral claim: it is hard to see how one could be more justified in bending to the coercion of another agent than to the coercion of the inanimate world.<sup>7</sup> However, as a claim about our intuitions about freedom, I think it is quite wrong. The distinction between how our fellow agents manipulate us, and how the inanimate world constrains our actions, is far from a mere matter of terminology; it lies at the heart of our ordinary conception of what it is to be autonomous.

I think that there is some intuitive plausibility to this: when we give examples of loss of autonomy it is common to cite a manipulating agent. But it is hard to disentangle this from our intuition about free will as the compatibilists have tried to characterize it.<sup>8</sup> So I shall approach the issue from a very different direction, one that I think will enable us to keep the two issues apart. My interest will be in our motivational structures.

What are our fundamental motivations? There are the obvious physical ones: food, shelter, sex. But in addition there is a set of fundamental social motivations, fundamental in the sense that they are almost universal, and that we generally cannot flourish if we fail to achieve them.<sup>9</sup> Although there is much debate over exactly how they should be classified, there is a fairly widespread agreement.<sup>10</sup> The first is a desire

---

<sup>6</sup> 'Coercion and Moral Responsibility', at pp. 45–6.

<sup>7</sup> In similar vein, Gary Watson argues that it is hard to justify the restriction of the legal defence of duress to cases in which one is coerced by another agent (rather than some other feature of the world). 'Excusing Addiction' in *Agency and Answerability* (Oxford: Clarendon Press 2004) 318–50, at p. 344.

<sup>8</sup> Compare Dennett's characterization of the Bogeyman as one of the metaphors used (in his view misleadingly) to motivate the free will problem. *Elbow Room* (Cambridge: MIT Press, 1984) pp. 7–10.

<sup>9</sup> They are thus plausibly *needs* as well as desires. For a fuller account of the ways in which desires or needs can be fundamental see R. Baumeister and M. Leary, 'The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation', *Psychological Bulletin* 117 (1995) 497–529.

<sup>10</sup> For an overview see the Introduction to E. T. Higgins and A. Kruglanski (eds) *Motivational Science: Social and Personality Perspectives*, (Philadelphia: Psychology Press, 2000)

for *social acceptance*.<sup>11</sup> The second is a desire for *control*: we become depressed and apathetic when we find that we cannot control our environment, either because it is uncontrollable, or because we lack the necessary competence.<sup>12</sup> The third, which is the one of relevance for us, is a desire for *self-determination*.

The idea of self-determination has been articulated and explored in the work of Edward Deci and Richard Ryan. They write:

Some intentional behaviors, we suggest, are initiated and regulated through choice as an expression of oneself, whereas other intentional behaviors are pressured and coerced by intrapsychic and environmental force and thus do not represent true choice. The former behaviors are characterized by autonomous initiation and regulation and are referred to as self-determined; the latter behaviors are characterized by heteronomous initiation and regulation and are referred to as controlled.<sup>13</sup>

True choice, they go on to say, applies only to actions that involve “an inner endorsement of one’s actions, the sense that they emanate from oneself and are one’s own”. Philosophical readers will be immediately reminded of Frankfurt.<sup>14</sup> But the idea isn’t quite the same. Frankfurt is concerned with the issue of what it is to endorse, or better, to identify with, one’s *desires*. Deci and Ryan are concerned with the issue of endorsement of one’s *actions*. The two are importantly different; and here we start to see the connections with our topic of coercion. When I am coerced into some wicked act to protect the life of my child, the desire which moves me—the desire for the well being of my child—is certainly one that I endorse; the wicked action is not. It is only in Deci and Ryan’s sense that the action is not my own.

It is tempting to try to give a philosophical analysis of the rather vague notion that Deci and Ryan are after. But care needs to be taken, for the notion is driven by its empirical explanatory value, rather than by an attempt to articulate an existing ordinary language concept. Its explanatory value turns out to be great. Factors that threaten self-determination, such as rewards, threats, deadlines, and even simple evaluation and

---

<sup>11</sup> See Baumeister and Leary, *op cit*. This is what underpins the mechanism of what Pettit and Brennan call ‘the intangible hand’. See ‘Hands Invisible and Intangible’, *Synthese* 94 (1993) 191–225.

<sup>12</sup> The former is the basis of Seligman’s notion of Learned Helplessness; the latter, of Bandura’s notion of Self-Efficacy.

<sup>13</sup> E. Deci and R. Ryan ‘The Support of Autonomy and the Control of Behavior’, *Journal of Personality and Social Psychology* 55 (1987) 1024–37, at p. 1024. In a footnote to this passage they remark that the distinction should be understood as marking the ends of a continuum, rather than a sharp break. Although they talk here of control coming from environmental force, in fact this seems to be entirely other people.

<sup>14</sup> Especially of his later writing where, as Scanlon notes, the stress moves to the issue of identification, and away from that of moral responsibility. See ‘Identification and Externality’ in *The Importance of What We Care About*, and ‘The Faintest Passion’ in *Necessity, Volition and Love* (Cambridge: Cambridge University Press, 1999). Scanlon’s discussion is in his ‘Reasons and Passions’ in S. Buss and L. Overton (eds.) *Contours of Agency* (Cambridge: MIT Press 2002), pp. 165–83. For a straightforward presentation of exactly what Frankfurt means by identification (i.e. acceptance, rather than endorsement or caring about) see his reply to Watson in that volume at pp. 160–1.

surveillance, tend to undermine *intrinsic* motivation, i.e. the motivation to go on with the activity even when the pressures are removed. They also undermine interest, enjoyment, and creativity. They tend to lower the trustingness of those exposed to them, increase their aggression, and make them, in turn, more controlling. Even their health suffers.<sup>15</sup>

Intrapersonal factors appear to have similar effects (though this is less well explored). So, for instance, being told that success in a task is an indicator of high IQ makes university students less motivated to go on with that task once they are given the chance to drop it. Even the self-surveillance provided by the presence of a mirror lowers intrinsic motivation. Deci and Ryan hypothesize that people can pressure themselves in much the same way as they can be pressured by others, and with much the same consequences. They go so far as to conclude

When behavior is prompted by thoughts such as “I have to ...” or “I should ...” (what we call internally controlling events), the behavior is theorized to be less self-determined than when it is characterized by more autonomy-related thoughts such as “I’d find it valuable to ...” or “I’d be interested in ...” Accordingly, we predict that the qualities associated with external controlling events and with external autonomy-supporting events will also be associated with their intrapsychic counterparts.<sup>16</sup>

Here I do voice some philosophical scepticism. It is very hard to believe that when things matter so much to us that we feel we *have to* act on them—whether this be Luther’s religious commitments or the commitment that a parent has to a child—we will feel a lack of self-determination.<sup>17</sup> It is not, I suspect, obligation in itself that undermines self-determination; it is rather our belief about the *source* of that obligation.

Clearly much of what is at issue in both the interpersonal and the intrapersonal case has to do with manipulation: we do not like being pushed about, not by others, and not even by our own demanding selves. Behaviour is perceived as more of a threat to self-determination when it is perceived as more manipulative. In this dimension, all rewards are not the same. Those offered for the performance of specific tasks within an experiment do more to undermine intrinsic motivation than those offered for simple participation: subjects perceive the experimenters as doing more to manipulate them.<sup>18</sup>

This has been a fairly lengthy detour. But I hope that it has succeeded in showing that the idea of self-determination is central to our idea of autonomy, which is one of our ideas of freedom. It is not the idea that compatibilists are trying to get at. But it is, I think, what is compromised in cases of coercion. Indeed, ‘compromised’ is far too

---

<sup>15</sup> For a recent set of studies see E. Deci and R. Ryan, *Handbook of Self-Determination Research* (Rochester: University of Rochester Press, 2002); for a popular presentation of the approach, see E. Deci, *Why We Do What We Do*, (Harmondsworth: Penguin, 1996)

<sup>16</sup> Deci and Ryan, ‘The Support of Autonomy’ *op. cit.*

<sup>17</sup> Frankfurt provides a lengthy discussion of this phenomenon, which he terms ‘volitional necessity’. See for instance ‘Rationality and the Unthinkable’ in *The Importance of What We Care About* and ‘On the Necessity of Ideals’ in *Necessity, Volition and Love*.

<sup>18</sup> E. Deci and R. Ryan, *Intrinsic Motivation and Self-Determination in Human Behavior* (New York: Plenum Press, 1985) pp. 72ff.

weak a word. For coercion provides as radical a subversion of self-determination as one can imagine. Unlike simple incentives, it typically involves a complete disruption of one's own plans in a most unwelcome way.<sup>19</sup> Coercion doesn't provide another option to be considered alongside what one is already doing; if it is successful, it requires us to abandon what we are doing. And even if we resist it, it will typically hijack our thoughts and our emotional energies. Of course, natural disasters can have many of these features too: they can disrupt our plans in most unwelcome ways. What is special about coercion is that another agent is seeking to manipulate us by deliberately employing such disruption. That is what makes the loss of self-determination so egregious.

My contention is that it is the manipulative feature of coercion that explains our metaphysical intuitions about its effect on freedom. But this does not explain our moral intuitions. First note that in general whilst we think that loss of self-determination is a bad thing, we don't think that it has an effect on moral responsibility. A person who is in jail has their self-determination greatly reduced, but we don't think that they are thereby relieved of any moral responsibility. Secondly, and more specifically, it appears that there are cases in which we think that moral responsibility is lessened in much the same way as it is in cases of coercion, even though there is no loss of self-determination.

Imagine the kind of case that gets endlessly discussed in introductory classes on consequentialism: one can only save one's dying child if one robs a pharmacy to get the drugs they need. Put aside the question of whether the robbery might be morally justified, and just consider the moral responsibility of someone who went ahead with it. We would surely think that their responsibility was much reduced (compared, say, with someone who committed a similar robbery to fund their summer vacation). It seems, moreover, that the person who robs the pharmacy to save their child has a similar kind of moral defence to the person who robs a pharmacy because they have been coerced into doing so by somebody who will otherwise destroy the drugs their child needs. I do not deny that there are differences between the two cases; we shall return to some of them later. Yet both seem to be of a moral piece, despite the fact that the second involves a violation of self-determination that the first does not.<sup>20</sup>

---

<sup>19</sup> Robert Nozick stresses the idea that coercion presents an unwelcome choice in 'Coercion', S. Morgenbesser, P. Suppes and M. White (eds.) *Philosophy, Science and Method* (New York: St Martin's Press 1969) pp. 440–72, at pp. 458ff.

<sup>20</sup> Although it remains controversial, English law now fairly clearly recognizes a defence of 'duress of circumstances', as part of a general defence of necessity. See J. Smith and B. Hogan *Criminal Law*, Ninth Edition (London: Butterworths, 1999) pp. 242–3, 245–52. They comment that whether the defendant is driven by a threat from an aggressor, or by natural circumstances, 'his moral culpability, or lack of it, is exactly the same' p. 23. The status of a parallel 'defense of situational duress' under US law is less clear; but the Model Penal Code proposes a general defence of necessity at Section 3.02.

## MORAL INTUITIONS

How then should we explain the effects that coercion has on our moral responsibilities? There are three obvious forms that such an explanation might have. At a first pass we might say:

- (i) a coerced action is not a free action at all; or
- (ii) a coerced action is a free action, but one that the agent was justified in performing; or
- (iii) a free action, one that the agent was not justified in performing, but one that we nevertheless excuse.

When I talk here of actions that are not free, I don't merely mean those in which the agent is manipulated. I mean what compatibilists have traditionally meant by lack of freedom. To claim that an action is not free is to say that in some important sense it was not the agent's action at all: perhaps the movement was involuntary, or the agent was pushed, or someone had taken control of their brain. Once we realize that an action falls into this class, the question of moral justification or excuse just doesn't arise.

To take the first explanation is thus to assimilate coerced actions to these sorts of action. This, I take it, is Pettit's project. It is also a path that some more conventional compatibilists have taken. Most notably Frankfurt has argued that coercion only occurs when 'the victim's desire or motive to avoid the penalty with which he is threatened is ... so powerful that he cannot prevent it from leading him to submit to the threat'.<sup>21</sup> A threat that doesn't move the victim in this way is mere duress; submission is not excusable.

I do not find this approach compelling. In the first place there are difficulties with the details of Frankfurt's account. What is it to be unable to prevent a desire from leading us to action? A natural way to understand this, and one which Frankfurt's comments suggest, is in terms of his own higher-order account of freedom: one is unable to resist a desire iff, were one to desire to resist it, one would not be able to. But that raises a host of problems that beset any conditional analysis: what, for instance, if one would indeed be able to resist if one were to form the desire to resist, but one was unable to form such a desire?

There are replies that might be made on Frankfurt's behalf to such worries; or we might try to salvage the general approach by proposing another account of what makes a desire irresistible. But I think that there are more fundamental problems with any account that tries to deny that coerced actions are free in the kind of way that compatibilists normally understand freedom. To begin with, note that coercion doesn't

---

<sup>21</sup> Frankfurt, 'Coercion and Moral Responsibility', p. 39. This is only a necessary condition for it to be coercion. In addition, the agent must have a desire or inclination to resist the desire, *ibid.*, p. 41

remove moral responsibility in all circumstances. This is reflected in the legal defence of duress. Under English law the situation is clear: duress can only be a defence if the threat is one of death or serious personal injury; and it is never a defence for murder.<sup>22</sup> In Blackstone's words: "A man ought rather to die himself than escape by the murder of an innocent". I take it that Blackstone is here making a moral claim, as well as a legal one. Phrased in such uncompromising terms that claim is controversial.<sup>23</sup> However even those who would reject it would normally accept some kind of principle of proportionality like that included in the American Model Penal Code: for duress to be a defence, the harm threatened against the defendant must be greater than that which the defendant's action can be expected to cause. Killing an innocent is hard to excuse, since it is so hard to see how there could be any circumstances in which the cost of resisting the coercion would be higher.<sup>24</sup>

But this suggests that we expect a certain degree of self-control of agents who yield to coercion. We expect them to be able to assess the gravity of the threat relative to that of the act that they are being pressured to perform. This does not suggest a picture of agents driven by irresistible desires. The point is brought out further by the fact that there is no subjective requirement in the law that a victim of duress should lose their self-control—in contrast to a provocation defence which requires that they do. One can still rely on a defence of duress if one's response has been as calculated as can be.

All of this in turn suggests that the right way of understanding coercion's moral status in terms of either the second or third models. A coerced action is a free action, but it is either justifiable or excusable. To suggest this is in no way original; this has been the approach that most theorists have taken. The main debate has concerned which of the two models is correct: whether the coerced person does a justifiable thing in bowing to the coercion, or an unjustifiable thing that we nevertheless excuse because no reasonable person could be expected to do otherwise.<sup>25</sup> I do not propose to try to resolve this question here. In fact it is unclear to me whether the two explanations are fully distinct. Very often we are concerned not with whether an action is justifiable *simpliciter*, but whether it is more or less justifiable: justifiability functions more as a scale than as an absolute threshold. Then we can unproblematically say that a coerced action

---

<sup>22</sup> Terminological point: the standard legal term is *duress*. In English law there is a defence of *coercion* but it is, oddly enough, normally restricted to crimes committed by women in the presence of their husbands.

<sup>23</sup> The idea that there can be no defence of duress for homicide is not a feature of the Model Penal Code. It remains, however, a feature of English law, having been recently endorsed by the House of Lords in *Howe*. For discussion see Smith and Hogan, *op. cit.*, pp. 231–44.

<sup>24</sup> *Ibid.* p. 252.

<sup>25</sup> For some recent discussions see J. Dressler, 'Exegesis of the Law of Duress: Justifying the Excuse and Searching for its Proper Limits', *Southern California Law Review* 62 (1989), 331–89; D. Kahan and M. Nussbaum, 'Two Concepts of Emotion in Criminal Law' *Columbia Law Review* 96 (1996) pp. 269–374; G. Watson, 'Excusing Addiction'; P. Westen and J. Mangiafico 'The Criminal Defense of Duress: A Justification, Not an Excuse—And Why it Matters' *Buffalo Law Review* 6 (2004) 833–950.

is typically more justifiable than a similar action that is not coerced, and frequently sufficiently justifiable to be excusable.<sup>26</sup>

What matters for my purposes here is independent of this debate though. For whether we understand coercion as providing justification or excuse, we will still think of coerced actions as free, at least as the compatibilists mean that. And in so far as there is a sense in which such actions are not free—in so far as we think of the coercion as removing self-determination—this is a feature that has no direct bearing on their moral status.

There remains, however, a puzzle, that anyone proposing an excuse or justification account should address. The puzzle concerns the relation between coercion and bribery. If I threaten to take away something which you already have if you do not do as I say, that is coercion. But if I offer to give you something that you do not already have if you will do as I say, that is bribery. Put this way the two do not sound very different; indeed there are going to be plenty of cases that come between them (I have promised you something which I threaten to withhold if you do not do as I say). Yet we do not generally think that the presence of a bribe removes the agent's moral responsibility (unless, that is, the agent is in such a state that the bribe removes their ability to think: offering cocaine to an addict for instance); in fact it will often make things worse. What is it about coercion that makes it so different?

There are various features that cases of coercion typically have that cases of bribery typically lack. As we have seen, typically people do not welcome the options presented by coercion, whereas they often welcome those presented by a bribe. And typically those who are coerced are dependent upon the coercer for something that they need, a dependence that the coercer exploits.<sup>27</sup> But these are not essential differences. We do not always welcome a bribe. Knowing that we are weak and are likely to succumb, it makes perfect sense not to want to be led into temptation. Sometimes too those who offer a bribe exploit a need that only they are in a position to meet. When we have a case of bribery that meets all of these conditions—a *coercive offer* as Frankfurt terms it—is it still obvious that succumbing to bribery is worse than succumbing to coercion?

I think that it is far from obvious. Consider variants of the case described above involving the parent whose child needs medication. In the coercive case the coercer steals the medication and refuses to return it unless the parent obeys. In the bribery case the parent does not yet have the medication, and the briber offers it if he will obey. As I said above, I think that the two cases are very much of a piece. However, I suspect that we are slightly more prepared to forgive the parent who succumbs to the coercion than the parent who succumbs to the bribe. This is especially so if we think of the case as involving excuse rather than justification. Suppose that what the parent does to get the medication involves something terrible—the killing of an innocent, say—so that we

---

<sup>26</sup> The issue is further complicated by whether we take justification to be agent neutral or agent relative; see Kahan and Nussbaum, *op. cit.*

<sup>27</sup> See Frankfurt 'Coercion and Moral Responsibility', p. 33

should never think it justified. We might, nevertheless, excuse the act; but we should do so more readily in the case of coercion than bribe.

The heart of the explanation of the difference surely lies in the fact that we think the coerced parent more deeply harmed, and so more readily excused in his response. Of course the briber is wrong to exploit his position; he should give the medication freely. But the coercer does an additional wrong in taking away the medication in the first place. There are two factors here: the wrongness of the theft, and the difference between losing what one already has, and failing to gain what one might have. We can control for the former by imagining a third case. Suppose that the parent has lost the medication. The coercer discovers where it is, but refuses to tell unless the parent obeys. We still, I think, would be more ready to excuse a parent coerced in this way than one who succumbed to the bribe.

I suspect that the explanation for this stems from the fact that we just do have a different attitude towards the loss of that which we already have, as opposed to our failure to gain that which we do not yet have. (The contrast can involve a possession, or something which is not properly thought of in this way, like peace of mind, or good health, or a child.) This attitude is manifested in another well-documented psychological phenomenon: *the endowment effect*. We value things more once we possess them.<sup>28</sup> The experiments that illustrate this effect often involve such patent irrationality that it is easy to think of it as a foible to which no importance should be attached. Thus, asked what they would pay for a Cornell University mug, the average subject offered around \$2.50; but, once given the mug, they were not prepared to give it up for less than \$5.<sup>29</sup> Yet, whether or not it is justifiable, this is no trivial tendency. We do think it far worse to deprive someone of their sight than to fail to restore the sight of someone already blind; and we tend to maintain that view even under critical reflection.

Perhaps this is merely an irrational tendency; certainly it has its irrational elements. Perhaps it reflects the importance we attach to the fulfilment of legitimate expectations. Perhaps it reflects some fundamental attachment to the status quo. I do not know where to start in pursuing this difficult question.<sup>30</sup> What matters here, since we are concerned with excuse and not with justification, is that we do tend to think in these ways; and hence we are prepared to offer greater excuse to those who think that they risk suffering a greater loss. I conclude that our different attitudes to bribery and to coercion can be explained within the excuse model.

---

<sup>28</sup> Closely related phenomena are *Loss Aversion* (losing something brings more cost than gaining it brings benefit) and the resulting *Status Quo Bias* (subjects have a strong tendency to keep with the status quo since losses consequent on change figure more prominently than gains). Loss Aversion is one of the underpinnings of Kahneman and Tversky's Prospect Theory; see their *Choices, Values and Frames*, (Cambridge: Cambridge University Press, 2000)

<sup>29</sup> D. Kahneman J. Knetsch, and R. Thaler 'Experimental Tests of the Endowment Effect and the Coase Theorem', *Journal of Political Economy* 98 (1990) 1325–48.

<sup>30</sup> The endowment effect plausibly has some work to do in explaining our different attitudes to doing and allowing. For a stimulating discussion see T. Horowitz, 'Philosophical Intuitions and Psychological Theory' *Ethics* 108 (1998) 367–85.

## FREEDOM AS DISCURSIVE CONTROL

I have been focussed so far on Pettit's negative arguments against the conventional compatibilist account of freedom. In this last section I turn to his positive proposal. Although Pettit is critical of the traditional compatibilist accounts, he does not deny that they provide necessary conditions for freedom. His argument is that that do not provide sufficient conditions. He thinks that they need to be supplemented with a further necessary condition: that of *discursive control*. My aim in this section is to give some consideration to this proposal, especially in the light of the kind of empirical considerations about human psychology that have figured so prominently up till now. Even if we are unconvinced by Pettit's argument against the standard compatibilist accounts, it is very plausible that those accounts do not tell the full story, especially if, as I think, the concept of freedom is a cluster concept. Does Pettit's notion of discursive control shed light on some aspect of the cluster?

Pettit's basic idea is this:

An agent will be a free person so far as they have the ability to discourse and they have the access to discourse that is provided within [discourse friendly] relationships"<sup>31</sup>

There are two parts to this. The first, the *ratiocinative*, concerns the abilities of the agent. In addition to the ability to deliberate, free agents must have the ability to discourse. The second, the *relational*, concerns the social situation of the agent: free agents must be able to enter into actual discourse, and to do this they must have others around them with whom they can have discourse-friendly relations. I start with the second.

### *The relational condition*

Pablo Neruda writes, in his paean to the Communist Party: "You have given me the freedom that the lone man lacks".<sup>32</sup> And certainly that may be right. There are many things that one can do with others that one cannot do on one's own. Moreover, the presence of others need not simply confer greater powers of execution: the ability to move heavy things, maintain a twenty-four hour vigil, outvote the opposition. Others can also help with the process of thinking. They can provide ideas, or force us to question and clarify our own, or enable us to maintain our commitments when we should otherwise despair.

Nonetheless it is hard to believe that actually available social relationships are necessary for freedom—does Crusoe really lose his freedom until Friday's arrival? And we might wonder about Pettit's claim that freedom, so conceived, is necessary for moral responsibility—if an interlocutor proposes actions so morally grotesque that everyone

---

<sup>31</sup> *A Theory of Freedom* p. 70.

<sup>32</sup> 'To my party' from *Canto General* (Berkeley: University of California Press, 1991) 298.

refuses to talk to him, do they thereby not only remove his freedom, but also absolve him of moral responsibility? Moreover, even the contention that the presence of others will bring us closer to the truth is far from obviously true. Whilst the jury theorem provides some *a priori* support for it, the empirical evidence is mixed. In some cases, especially those involving simple concrete questions, groups do better than individuals thinking alone. But in some cases they do worse. Part of this can be explained by the presence of a strong tendency to conformity or compromise.<sup>33</sup> This is especially marked when there is no demonstrable right answer: for instance, if the question posed concerns a difficult issue of value, or of which of various plausible actions should be taken. But even in the case of a simple empirical issue with a demonstrable right answer there is a strong tendency to convergence. One influential set of experiments by Solomon Asch has led to a large literature. Subjects were asked to judge the relative lengths of lines. The task was easy, and, asked on their own, 95% gave the right answer to each of a series of twelve or so tests. They were then placed in a group of six to eight others, confederates who had been briefed to answer correctly on two initial tests, but then to give wrong but unanimous answers on the remaining ten. The subjects could hear what the others said before they gave their own answers; and their answers were in turn heard by the others. In these circumstances only 24% gave the right answer in each of the tests where the others answered wrongly; 25% got more than two thirds of them wrong. A number of factors seem to be at work: the desire to win the approval of others (as shown by the fact that answers that cannot be heard by the others conform less); a desire not to seem deviant even to oneself; and the conviction that the others must be right.<sup>34</sup>

These factors are surely politically and morally important: they go some way to explaining how politics is possible in a world of opposed interests. Yet it is hard to see how they are important in an account of freedom. If anything they seem to involve a subordination of the capacities involved in rational control to wider social goals. Individuals willingly (or perhaps unknowingly) give up certain freedoms in belief and desire for the wider good. So let us turn from the issue of whether the free agents need to have actual social relationships to the issue of whether they need have the ability to engage in them.

---

<sup>33</sup> I will not discuss the literature on the tendency to compromise, though there is one experiment whose outcome I feel duty bound to report: academics showed themselves far more likely to agree to see a student for a single twenty minute meeting if they had previously refused the (surely, one hopes, unreasonable) request to meet weekly for two hours for the rest of the semester. See H. Harari, D. Mohr and K. Hosey, 'Faculty Helpfulness to Students: A Comparison of Compliance Techniques' *Personality and Social Psychology Bulletin* 6 (1980) 373–7. For a review of the literature in this area see R. Cialdini and M. Trost, 'Social Influence: Social Norms, Conformity and Compliance' *The Handbook of Social Psychology* Vol. II, pp. 151–92 at pp. 177ff.

<sup>34</sup> For review see Cialdini and Trost, *op. cit.*, pp. 162 ff. There are many other factors that affect the way that dialogue in fact develops, including size of the group, gender, familiarity, and whether we are acting for ourselves or for others. For a general overview see J. Levine and R. Moreland 'Small Groups' in *The Handbook of Social Psychology*, Vol. II, pp. 415–69

*The ratiocinative condition*

The idea here is that in order to enter into discourse an agent must have various abilities; and it is the possession of these abilities, rather than the process of discourse itself, that is necessary for freedom.<sup>35</sup> What abilities does an agent need to enter into discourse over and above the ability to deliberate? They must, of course, have a common language with those around them, and the ability to hear and be heard (or to read and be read, or whatever). Those are certainly crucial abilities, although, given the social dimensions of knowledge, they are plausibly required even for rational deliberation. What else is needed that is distinctive to discourse? Pettit is not altogether clear about this, but there is one plausible condition. On Pettit's conception a discourse-friendly relationship is one in which agents can 'reason together'.<sup>36</sup> To do this, discoursing agents must have the ability to justify their beliefs and their actions to each other; it is only through such justification that legitimate influence can take place.

If this is right, then the ratiocinative requirement certainly does add a further condition: to be free, agents must be able to justify what they think and do. But this is a highly controversial condition. There is good reason to think that in exercising many capacities we simply don't know what it is that we are responding to, so we are in no position to justify either our beliefs or our actions. Chicken-sexing is the standard philosophical example, but a bad one: it appears that most chicken sexers do know which features they are picking up on when they form their beliefs.<sup>37</sup> Human-sexing is a better example, at least when we have just the face to go on. Most of us are extremely fast and accurate at telling someone's sex by looking at their face. How do we do it? It turns out that we are sensitive to many cues: for instance, the shape of the nose, especially the bridge, is crucially important. Similarly we are good at telling someone's age from their face, and here we go mainly, not by wrinkles as we might expect, but by overall shape and colour distribution.<sup>38</sup> In both cases we have little clue about what we are doing until we are told. Many other capacities, including learned capacities, are similar. Gary Klein gives a large number of examples involving soldiers, sailors, fire-fighters, doctors and neo-natal nurses.<sup>39</sup> In many cases we are simply unable to justify our judgements and decisions. Experienced fire fighters have a gut feeling when a burning building is particularly dangerous; experienced nurses just know when a

---

<sup>35</sup> Compare the distinction that T. M. Scanlon makes between his account of contractarianism, in which actual agreement with others is not required, and the position that he attributes to Habermas, in which it is. See *What We Owe to Each Other* (Cambridge: Harvard University Press, 1998) p. 393, n. 5.

<sup>36</sup> Pettit, *ibid.* pp. 69ff.

<sup>37</sup> I. Biederman & M. Shiffrar, 'Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual learning task', *Journal of Experimental Psychology: Human Learning, Memory, and Cognition*, 13, (1987) 640-645.

<sup>38</sup> For a nice presentation on both sex and age discrimination see V. Bruce and A Young, *In the Eye of the Beholder: The Science of Face Perception* (Oxford: Oxford University Press 1998)

<sup>39</sup> *Sources of Power* (Cambridge: MIT Press, 1998)

premature baby has started to develop an infection. Of course there are cues that these people are picking up on; they are not doing it by magic. But they do not know what those cues are. Moreover, often the attempt to justify them, or to make them using publicly justifiable criteria, actually corrupts our judgment.<sup>40</sup>

One response to this line of worry is to stress that we are dealing with an ideal: in fact people are unable to justify their beliefs and decisions, but ideally they should. But for this line to work we need some reason to think that it is an ideal, and none has been offered. A second, more plausible, response is to say that the kind of judgements we are considering are justified, but that they are justified not by citing the evidence upon which they are based, but by citing the experience and past performance of the agents who make them. The judgments of the experienced fire-fighter and experienced neo-natal nurse are justified by pointing to the skills they have learned that are manifested in their past successes. That response seems exactly right for these sorts of cases, cases in which the agent has expertise in discerning some hard-discerned empirical fact. But it is less plausible when it is some difficult moral problem that is at issue. At least when we are dealing with adult agents, we are far less happy with the idea that a moral expert can justify a judgment solely on the basis of their track-record. Perhaps there are some thick moral concepts for whose application we are rightly prepared to defer to those who have shown they can apply them without being able to justify their application: being sexist for instance. But it is far less plausible that we would be accept a relatively abstract moral principle—‘Sexism is wrong’—to be justified on the basis of a testimony of a competent moral judge.<sup>41</sup>

However, rather than being an objection to Pettit’s account, this points to a defence. For the idea is perhaps that in presenting *moral* conclusions we need to be able to give justifications. We do not need to be able to justify every application of a thick moral concept along the way. But we do need to be able to justify the more abstract principles upon which any conclusion rests. The idea might be filled out in various ways. We might, for instance, think that justification affects our cognitive or epistemic

---

<sup>40</sup> See also Wilson and Schooler, ‘Thinking too much: introspection can reduce the quality of preferences and decisions’ *Journal of Personality and Social Psychology*, 60 (1991). Ordinary subjects, asked to rank samples of strawberry jam in terms of quality, gave rankings that were close to those of expert tasters, until they were told to use explicit criteria, at which point their rankings diverged substantially. Klein gives many examples of cases in which subjects do far worse when asked to make decisions by running through an explicit checklist of factors rather than on the basis of a single overall judgment. The distorting effects of the checklist approach will be familiar to those who have been instructed, by various regulatory bodies, to assign marks to undergraduate essays by summing the marks assigned to supposedly relevant factors: understanding, clarity, structure, breadth of reading, originality etc.

<sup>41</sup> The example comes from Karen Jones, ‘Second-hand Moral Knowledge’ *Journal of Philosophy* 96 (1999) 55–78. I think that she is clearly right that, in the application of thick moral concepts like ‘sexist’ we should be prepared to defer to the testimony of those who are more skilled at applying them; they do, after all, contain a considerable descriptive component. What I question is whether the same is true of more abstract principles (an issue which Jones raises, but on which she does not come to a conclusion).

standing with respect to a principle: a person must be able to justify a moral principle in order to understand it, or in order to know it. That strikes me as rather implausible version of the requirement. We would need to know why the power of testimony to confer knowledge was so much more restricted here than elsewhere. A more plausible approach is to think that the requirement is itself a moral one: whilst we might be able to gain knowledge of moral principles entirely on another's testimony, we have a moral obligation to understand the basis of that knowledge.

I am not sure what to think of such a putative moral requirement. I am confident that there was a time when it would have been broadly rejected: a time when accepting God's testimony, or that of His representatives on Earth, without needing to understand the reasoning, was the paradigmatic moral position. But it does seem far more plausible that it is a component of Enlightenment moral thinking.<sup>42</sup> Yet it surely stands in need of some justification. Even if it can be purged of the threat of regress which it obviously invites (Can *all* of our moral principles be justified?), we need to be assured that it does not rest on the same kind of foundationalist instincts that have been successfully challenged elsewhere.

Obviously this is a large issue, one that cannot be pursued here. Let me conclude instead by returning to the subject at issue, that of freedom. I have developed the idea of the ratiocinative condition as requiring that we be able to justify our beliefs and decisions; and I have suggested that if this idea has any plausibility it is as a moral requirement that we be able to justify our moral reasoning. How does this tie in to the issue of freedom? Pettit's proposal is that meeting the requirement is a necessary condition on being free. I think that there is some plausibility to the idea that being free requires a kind of moral ability. But I am doubtful that Pettit would accept it. For it takes us back to the idea of freedom as a kind of rational control, one of the targets of Pettit's criticisms.<sup>43</sup> Perhaps then I have misinterpreted the ratiocinative condition; at the least I hope that I have shown where one natural development leads.

*Department of Linguistics and Philosophy*

---

<sup>42</sup> As I read it, it is one of the elements of the position presented in Kant's *What is Enlightenment?* in M Gregor (ed.) *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy* (Cambridge, Cambridge University Press, 1996) 11–22. The thought there is not simply that one need not be afraid to think things through for oneself, but that there is a moral requirement to do so.

<sup>43</sup> The most striking presentation of this version of the idea is in Susan Wolf's paper 'Sanity and the Metaphysics of Responsibility', in F. Schoeman (ed.) *Responsibility, Character and the Emotions* (Cambridge: Cambridge University Press, 1988) 46–62. There though the focus is on the idea that certain moral capacities are necessary for moral responsibility, rather than that they are need for freedom itself. Note too that Wolf does not endorse any particular moral requirement that one be able to justify one's moral reasoning.

*MIT*  
*bolton @ mit.edu*