

MIT Open Access Articles

Determinism, Self-Efficacy, and the Phenomenology of Free Will

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Holton, Richard. "Determinism, Self-Efficacy, and the Phenomenology of Free Will." *Inquiry: An Interdisciplinary Journal of Philosophy* 52.4 (2009): 412-428.

As Published: <http://dx.doi.org/10.1080/00201740903087383>

Publisher: Routledge

Persistent URL: <http://hdl.handle.net/1721.1/50264>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



DETERMINISM, SELF-EFFICACY, AND THE PHENOMENOLOGY OF FREE WILL¹

RICHARD HOLTON, MIT

I. DETERMINISM AND MORAL MOTIVATION

Most of us, certainly most of us working in universities, like to think that the spread of knowledge will be beneficial. Diffuse knowledge, goes this pleasingly democratic line, and people will be better off. But there is another strain of thought, going back at least to Plato, that takes the contrary course, arguing that widespread knowledge of the truth will be damaging, and so insisting on the need for the Noble Lie.

Some recent work in social psychology has lent support to this latter approach, apparently showing that getting subjects to reject the existence of free will or to believe in the truth of determinism makes them less likely to behave morally. In a pair of studies Vohs and Schooler found that getting subjects to read some passages arguing that free will is an illusion subsequently made them more likely to cheat in a test.² In further studies, Baumeister, Masicampo and DeWall found that reading passages propounding determinism increased subjects' tendency to behave aggressively towards others (serving hot salsa to those who had said they hated it), and decreased their tendency to say that would behave helpfully in various situations.³

Assuming that these results are real, how are we to explain them? One obvious explanation will occur to philosophers: if the subjects believe that free will is necessary for moral responsibility, then, given minimal rationality, undermining their belief in free will should be enough to undermine their belief that they are morally responsible. And if they no longer think that they are morally responsible, immoral behaviour will follow.

Supporters of this interpretation might then point to a set of recent studies that seem to show that most people do think that moral responsibility is absent in a deterministic world, at least so long as they think sufficiently abstractly.⁴ If determinism is in fact true, the conclusion is a rather depressing one: we should either cultivate a belief in free will as a necessary illusion for moral behaviour; or else we need to embark on a probably fruitless attempt to convince the masses of the truth of compatibilism.

I want to suggest an alternative interpretation. Although it is true that some studies do suggest that most people believe that there is no moral responsibility in a deterministic world, it appears that this finding is highly sensitive to how the deterministic world is described. A recent study by Eddy Nahmias and colleagues has

¹ In addition to the Oslo Conference on Liking and Wanting, versions of this paper have been given at Florida State, Oxford, Edinburgh and Paris. I am very grateful to comments from the audiences there, and in particular to Kent Berridge, Elizabeth Fricker, Jennifer Hornsby, Joshua Knobe, Rae Langton, and Nick Shea.

² K. Vohs and J. Schooler, 'The Value of Believing in Free Will' ms.

³ R Baumeister, E. Masicampo and C.N. DeWall, 'Prosocial Benefits of Feeling Free' ms.

⁴ See, for instance, S. Nichols and J. Knobe, 'Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions', *Noûs* 41 (2007), 663–685. For the caveat about the need for abstract thinking see Nahmias *et al.*, 'Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility', *Philosophical Psychology* 18 (2005) 561–584.

found that what subjects really find inimical to moral responsibility is mechanism: the idea that we are being pushed along by happenings at the molecular level.⁵ And this finding is even more marked when they are asked, not about moral responsibility, but about freedom. This opens the door to the possibility that the apparent commitment to incompatibilism stems from a misunderstanding of the true nature of determinism, one that sees it as more mechanistic than it need be. And on the basis of this we might try for an alternative explanation of the Vohs and Baumeister results: perhaps the deterministic texts are encouraging a mechanistic view of the way the world works, and the moral demotivation stems from the subject's belief that that is incompatible with moral responsibility.

I think that this interpretation is a step in the right direction. But it still involves a highly theoretical bit of reasoning from the subjects: they have to think about the nature of morality and of the conditions that it requires. Mightn't the explanation be rather more primitive than that? Sticking with the thought that stories of determinism encourage an overly mechanistic view of the world, mightn't it be the case that this view is itself morally demotivating? Perhaps reading the deterministic texts tends to make agents think of themselves as powerless. This can pick up on a very normal distinction that people ordinarily make, between the things we can change, and the things we can't. So let me start by looking to the nature of that distinction.

WHAT YOU CAN CHANGE AND WHAT YOU CAN'T

Luckily there is an obvious place to begin, with Martin Seligman's admirably straightforward book *What you can change and what you can't*.⁶ Seligman describes the book as a guide to self-help guides. His focus is on the things that people typically want to change about themselves and on the extent to which, by deciding to do so and employing the right approach, they can succeed in doing so. It turns out that some things are relatively easy to change. If you suffer from panic attacks, then with application and the right kind of cognitive therapy, you should be able to get over them. Around 95% of sufferers do. Other conditions are more rarely shaken: if you are an alcoholic then there is around a 33% chance that you will succeed in giving up. For other the success rate is even lower. If you are overweight and want to slim by dieting, there is very little chance—around 5%—that you will much reduce your weight in the long run.

A number of things are interesting about this. Obviously one is that the relative likelihood of change is not what one would expect: we know that dieting is hard, but it is shocking that success is that uncommon. A second, more philosophically interesting, point is that Seligman is working with a notion of what can and can't be changed that is statistical or probabilistic. It is not that it is impossible to permanently lose weight by dieting, in the sense that no one ever does it. Five per cent do. It is rather that it is

⁵ E. Nahmias, D. J. Coates, and T. Kvaran, 'Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions', *Midwest Studies in Philosophy* 31 (2007) 214-42. The study found that people think that neurological determinism (everything we do is determined by the prior arrangement of our neurons) poses more of a threat to free will and to moral responsibility than does psychological determinism (everything we do is determined by our prior beliefs, desires and intentions). The authors interpret this as driven by a different reactions to more or less mechanistic conceptions of determinism. I think that that is very plausible, but the data could be interpreted rather differently. It could be seen as indicating an implicit belief in dualism.

⁶ M. Seligman, *What You Can Change and What You Can't* (New York: Knopf, 1993)

unusual to do so. Moreover, and here is a further philosophically interesting point, if we look in more detail at why that is, we find that it is because such change is difficult: it requires a tremendous amount of effort. Those who try to reduce their weight below a certain level typically become obsessed by food. They can think of little else. And in the end this incessant intrusion wears them down.⁷

Philosophers have historically been little concerned with such probabilistic measures. In looking at what an agent can or cannot do, the conception is typically absolute. Either one can lose weight or one can't. And this is all the more marked when we look to discussions of freedom of the will. Typically this is thought to be a property of agents, rather than of particular actions. Beings either have free will or lack it: the classic question is whether human beings, or indeed any possible beings, fall into the first or the second class. Even when freedom is relativized to actions—when philosophers talk of an agent being free *to do a certain thing*—the conception is still typically absolute: the thought is that the agent either is or is not free to do it. In contrast Seligman explicitly takes a graded approach to freedom itself: an agent is more or less free to perform a certain action depending on how difficult that will be. At the limits perhaps there are things that no one will do however hard they try; and others where no one who puts their mind to it will fail. But the interesting cases fall in between.

I don't specially want to castigate philosophers for working in absolute terms; I think that this is a very natural way of thinking. Indeed it is reflected even in the title of Seligman's book: *What you can change and what you can't*, not the more accurate (though admittedly rather less punchy) *What is easy to change, and what isn't*. But more than this, I suspect that philosophers have a suspicion that there is no sense to be made of the idea of degree of effort, the idea that is needed in accounts like Seligman's.⁸ I think that findings like those summarized by Seligman, and other findings that I will discuss later, show that we have to make sense of it: it is the best explanation of why some things are harder to change than others.

Of course, motivation plays a role too: one is much more likely to succeed in a task that takes a lot of effort if one is well-motivated to do so. But we have no reason to think that differences in motivation are the major explanation of the differences in success, so I won't discuss that further. It also matters which techniques one uses to try to effect a change. Panic attacks are relatively easily cured using cognitive behavioral therapy, but do not respond well to psychoanalysis. Clearly this is something that we do need to factor in.

So let us give a rough characterization of a graded constraint on freedom as follows:

Given the techniques they have at their disposal, an agent is less free to F the harder it is to follow through on a decision to F.⁹

⁷ For an accessible survey of some of the empirical data on how hard it is to lose weight by dieting see Gina Kolata, *Rethinking Thin* (New York: Farrar, Straus and Giroux, 2007). Of course this is a controversial issue, upon which I don't want to rest too much. If you are sceptical, substitute one of the other things that are very hard to change, such as sexual orientation (for 'pure' homosexuals, not bisexuals).

⁸ See for example Gary Watson. In his excellent piece 'Skepticism about Weakness of Will' *Philosophical Review* 86 (1977), 316-339 Watson endorses an account based around the idea of degree of difficulty of resistance, but denies that there is any place for a notion of effort.

⁹ Don't put the definition in classic conditional form: 'if one were to decide to F one would ...'. We know from Frankfurt cases that that will fall foul of the conditional fallacy; moreover, it is quite unclear how to convert it to a graded form.

This isn't a definition of freedom; for a start, there may be other ways in which freedom could be lost or diminished. It rather provides the basis of a necessary condition: one is not very free if following through is very difficult. Clearly even in specifying a necessary condition the account it offers is far from reductive. It leaves open quite what the relevant kind of difficulty is (it is the kind of thing that is more present in sticking to a diet than in the regulation of breathing needed to avoid panic attacks), and what it is to follow through (we are talking about success in achieving the actions immediately under the control of an agent, not their more distant consequences). But it should be clear enough to be going on with.

There are a lot of philosophers to whom this won't be of much interest. For them the interesting question will be: are you free to make the decision to F in the first place? And the worry is that determinism shows that you are not. I don't want to deny that this is an interesting question; it is an instance of a radical skepticism which philosophy has always been good at pursuing. But just as Cartesian skepticism isn't the only interesting question in epistemology, so this radical skepticism isn't the only interesting issue surrounding freedom. I want to suggest that thinking about freedom in the way that Seligman does helps us understand more about where our intuitions of freedom come from; and that whilst this in turn doesn't answer all of the questions about moral responsibility—it doesn't help with radical skepticism—it does help to explain the issues about our moral motivation with which we started. I'll start with the source of our intuitions of freedom.

INTUITIONS OF FREEDOM

As I mentioned at the outset, with the right prompting, it is easy to elicit incompatibilist intuitions from people. But these intuitions aren't standardly coupled with a belief in determinism. Rather they are coupled with a belief that we have free will. Where does this belief come from?

Sometime it is said that we have a direct experience of freedom. But if freedom is really understood to be something that is incompatible with determinism, it is hard to know what such an experience would be like. What is it to experience one's action as not causally determined, or oneself as an uncaused cause? I have no idea how that could be the content of an experience. The complaint here is an old one: Anthony Collins objects to those who appeal to vulgar experience to support libertarian views, 'yet, inconsistently therewith, contradict the vulgar experience, by owning it to be an *intricate matter*, and treating it after an intricate matter'.¹⁰

What we need, I think, is to identify ways in which action is experienced as something that the agent instigates, rather than something that just happens to the agent as the result of the state that they were antecedently in. And it will be all the better if we can have some explanation of how the experience of that state could be somehow mistaken for an experience of being an uncaused cause.

I will suggest two complementary possibilities. The first I have written about elsewhere, so I shall outline it quickly. The second is my focus here.

¹⁰ *An Inquiry Concerning Human Liberty* (Second Edition 1717) p. 30.

FIRST EXPERIENCE OF FREEDOM: CHOICE

In an earlier paper I argued that one important source of our belief in free will comes from our experience of choice.¹¹ A quick and effective way to generate (or, more likely, to make salient) a conviction in undergraduates that they have free will is to get them to focus on what choice is like. Tell them to make an arbitrary choice, and then get them to act on it—to raise their left hand or their right, for instance—and they are, by and large, left with an unshakable conviction that their choice was a free one.

What is happening here? They have in the first instance an experience: an experience of doing something, namely making a choice. Quick on its heels comes a judgement, or a clutch of judgements: that they could have made either choice; or, more theoretically committed, that they could have made either choice compatibly with how they were prior to the choice; or more committed still, that they could have made either choice no matter how the whole world stood prior to the choice, and hence that they are, in that respect, unmoved movers.

As I have said, judgments like these last surely go well beyond the contents of the experience. How could one have experience of the rest of the world in a way that revealed that one's action was itself uncaused? But the experience itself is an experience of something. At its heart, I suggest, are two aspects. First, we have an experience that provides the basis for a belief in the efficacy of choice, by which I mean that, in cases where the question of what to do arises, choice is both necessary and sufficient for action—choose to raise your right hand, and you'll raise it, likewise with your left; fail to make either choice and you won't raise either.¹² Second, we have an experience of different choices being compatible with our prior beliefs, desires and intentions. Believing, desiring and intending as one does, one could either choose to raise one's left hand or one's right hand. In particular then, choice raises the possibility of *akrasia*, since one might judge that it was best to raise one's left hand, and yet still raise one's right.¹³

It is easy to see then how this experience could be mistaken for the experience of one's actions being undetermined: one is mistaking a local claim of nondetermination (one's action is not determined by one's prior beliefs, desires, intentions etc.) for a global one (one's action is not determined by anything). This could generate the incompatibilist beliefs. Alternatively, it could be that the incompatibilist beliefs come from an overly mechanistic understanding of what determinism is like. It might be understood to undermine the effectiveness of choice as I have characterized it. There is a common way of thinking about determinism according to which it leaves the agent completely passive. But before saying more about this, let me introduce what I take to be the second source of our intuitions about our freedom. For the mental activity needed to form an intention is not the only source of activity needed prior to action: there is also sometimes activity

¹¹ 'The Act of Choice', *Philosophers' Imprint* 6, (2006).

¹² The qualification 'when choice arises' is an important one. If much contemporary social psychology is right, most actions are performed automatically, with no need for choice.

¹³ On this approach it is an interesting question as to why we have choice. It can look like a liability, opening as it does the possibilities of *akrasia* and inaction. Wouldn't we do better as creatures whose actions are linked directly to what we judge best, circumventing any need for choice? Briefly, my answer was that, as cognitively limited creatures, we are frequently unable to make a judgment as to what is best. We need to be able to choose to act even in the absence of such judgements, especially if can be influenced by relevant factors, even if these are not perceptible at the level of judgement. For instance, it looks as though we can be sensitive to patterns in the world, so that our choices may be influenced by them, even though we form no conscious beliefs about them.

needed to turn the intention into action. And this is especially important for our concerns here, since whilst finding oneself unable to make a choice is pretty unusual (it happens perhaps in cases of Parkinson's, and depression, and certain forms of prefrontal damage), finding intentions too hard to maintain is all too familiar.

SECOND EXPERIENCE OF FREEDOM: MAINTAINING RESOLUTIONS

Intentions play various roles in our mental economies. As Michael Bratman has stressed, they can help us to curtail deliberation, and can enable us to achieve inter- and intra-personal coordination.¹⁴ In addition there is a special class of intentions that enable us to persist in an action even when we know that there will be strong inclinations to abandon it: inclinations stemming from temptation, fear, fatigue, or whatever. Call such intentions *resolutions*. Resolutions need to be particularly resistant to reconsideration, since it is in the nature of the states like temptation and fear that they tend not simply to provide strong contrary inclinations, but to corrupt our judgments as well.

Putting most intentions into action doesn't require much on the agent's part: if she has decided to stop off to buy milk after work, then typically all that is needed is that she keeps the intention in mind. Action will follow. But in the case of resolutions things are very different. As we know from personal experience, implementing a resolution is difficult. This is just what is shown by the evidence that Seligman cites. For clearly the decisions involved in the cases he considers are resolutions: resolutions to eat less, drink less, control one's fear and so on. And this is confirmed by a host of further empirical work.

Consider first the fact that the ability to abide by a resolution is affected by features that do not themselves seem to be desires or resolutions. Reformed alcoholics are far more likely to relapse if they are depressed, or anxious, or tired.¹⁵ Moreover states such as these affect one's ability to abide by *all* of one's resolutions: resolutions not to drink, not to smoke, to eat well, to exercise, to work hard, not to watch daytime television, or whatever. Now of course it is possible to explain this by saying that these states (depression, anxiety, fatigue etc.) systematically strengthen all of one's desires to drink, smoke, eat, etc., or weaken all of one's resolutions not to; but it is surely a more economical explanation to say that they affect one's ability to act in line with one's resolutions.¹⁶ For why else would there be such systematic effects?

Consider next the remarkable empirical literature on what is known as 'ego depletion'. It appears that will-power comes in limited amounts that can be used up: controlling oneself to eat radishes rather than the available chocolate-chip cookies in one

¹⁴ Michael Bratman, *Intentions, Plans and Practical Reasoning* (Cambridge: Harvard University Press, 1987)

¹⁵ R. Baumeister, T. Heatherton and D. Tice, *Losing Control* (San Diego: Academic Press, 1994) pp 151ff. The same is true of those who are dieting (*ibid.* pp. 184ff.), or trying to give up smoking (*ibid.* pp. 212ff.) or taking drugs (Muraven and Baumeister, 'Self-Regulation and Depletion of Limited Resources', p. 250).

¹⁶ Moreover, whilst bad moods make dieters want to eat more, they tend to have the opposite effect on those who are not on a diet. So it seems that it is the resolution being affected, not the desire. See Muraven and Baumeister 'Self-Regulation and Depletion of Limited Resources' p. 251.

experiment makes one less likely to persist in trying to solve puzzles in the next¹⁷; suppressing one's emotional responses to a film makes one less likely to persist, later on, in holding squeezed a handgrip exerciser.¹⁸

A final piece of evidence is that one can apparently develop one's faculty of will-power by repeated exercise. Again, the idea that one becomes virtuous by behaving virtuously is a commonplace one, stressed by Aristotle and by many who have followed him: 'From holding back from pleasures we become moderate, and also when we become moderate we are most capable of holding back from them'.¹⁹ Some research suggests that this might be right: subjects who undergo a regime of self-regulatory exercises—working on improving their posture for instance—show markedly less tendency to suffer ego-depletion.²⁰

What exactly is the nature of the effort involved? We can get some handle on it by considering some of the most extreme cases, namely those involving addiction. People sometimes speak as though addiction hi-jacks the body: as though one's intentional systems are simply by-passed. But whilst that may be true of some actions—certainly reflexes like startle, and perhaps, more interestingly, some actions performed in response to great fear²¹—addiction typically works through, rather than against, our intentional systems. There is good reason why this should be so. Meeting an addiction will typically require some real planning on the agent's part: finding a dealer, negotiating a price, avoid arrest. It isn't the kind of thing that could be achieved entirely by automatic actions.

It is the fact that addiction works through one's intentional system that makes resistance possible. What is it then that makes it so hard? Work by Robinson and Berridge has indicated that addiction involves, very roughly, a decoupling of wanting and liking. Addicts need not like the substances to which they are addicted: they need take no pleasure in getting them, nor even in the prospect of getting them. Instead, addiction typically involves a sensitization of some parts of the mesolimbic dopamine system, with the result that perception, or just contemplation, of the addictive substance provides a strong motivation to get it.²²

¹⁷ R. Baumeister, E. Bratslavsky, M. Muraven and D. Tice 'Ego-depletion: Is the Active Self a Limited Resource?' *Journal of Personality and Social Psychology* 74 (1998) pp. 1252–65. The puzzles were in fact insoluble.

¹⁸ Muraven, Tice and Baumeister, 'Self-control as a Limited Resource'.

¹⁹ *NE* 1104a34. Aristotle is here talking about how we develop the excellences. He does not explicitly say the same about the development of self-control though. He does say that lack of self-control can be cured, but he doesn't say how.

²⁰ M. Muraven, R. Baumeister and D. Tice, 'Longitudinal Improvement of Self-Regulation Through Practice: Building Self-Control Strength Through Repeated Exercise' *The Journal of Social Psychology* 139 (1999) pp. 446–57.

²¹ Geoffrey Keynes talks of his experiences as a doctor in the trenches in the First World War, finding his legs reluctant to take him towards the falling shells, despite his intention to go and tend to the wounded. Keynes, *The Gates of Memory* (Oxford: Clarendon Press, 1981) p. 137. Of course that involves lack of action, not action. It would have been odder if his legs had started walking of their own accord. Odder still if his body had embarked on a complex series of actions—walking to his horse, mounting it, riding back to his quarters, writing a letter to his commanding officer asking for compassionate leave or whatever.

²² For a survey of a large body of work see T. Robinson and K. Berridge, 'Addiction' *Annual Review of Psychology* 54 (2003) 25–53. As against rational choice accounts which understand addiction in terms of a desire to avoid the pains of withdrawal, Robinson and Berridge's account explains how the desire for the addictive substance can start before withdrawal sets in, and can continue long after it has passed. Indeed there is some evidence that the sensitization might be permanent.

Now it might seem that the absence of liking would make it easier to resist: after all, it means that one just has to struggle against the urge to take the substance, and not also against the motivation that would be provided by liking it.²³ But that is to misunderstand how human beings work. The point is that normally liking has its effect through the motivational system; increased liking leads one to want more. Conversely, if one likes something less, one is less motivated to pursue it. Once wanting and liking are decoupled that is no longer so: the motivational system goes on, despite changes that would normally act as a brake upon it.

It is plausible that much the same can be said about the decoupling effected between wanting and the more cognitive attitude of judging a thing good. Standardly if I want something—a clever device for peeling garlic say—and then discover it doesn't work, my want will simply evaporate. But in cases of addiction there is an almost complete disconnection here. Coming to judge that the addictive substance is bad for me has very little effect on motivation to get it.²⁴ Indeed, in so far as there is a tendency for motivation and judgments about what is good to line up in cases like addiction, it is plausible that the influence goes in the other direction. Knowing the way in which they will act, subjects rewrite their judgments about what is good to avoid cognitive dissonance.²⁵ Nonetheless, these judgments will have further causal impact—ego depletion shifts agents' choices about what to consume later as well as choices about what to consume now.²⁶

The kind of wanting involved in addiction does not appear fundamentally different to the kind of wanting involved in other more ordinary circumstances. It is strong, but it works in much the same sort of way: by capturing the attention, focusing though on what is desired, narrowing horizons. In cases of addiction this is extreme enough to count as craving; but this too can be a feature of non-addicted desire.²⁷ What is most saliently different is that, since it is decoupled from liking and from beliefs about what is best, it cannot be changed by changes in them. Impervious to undermining, it can only be directly resisted.

²³ For something like the picture that the different forces are additive, see David Velleman's 'What Happens When Someone Acts', *Mind*, 101, (1992), pp. 461-481.

²⁴ It might be argued that states disconnected in these ways should not count as desires; Scanlon, for instance, writes that 'what is generally called having a desire involves having a tendency to see something as a reason' (Scanlon, 1998) pp. 39. That strikes me as either wrong (if viewed as a description of our current practice of using the term), or as an unhelpful stipulation (if viewed as an instruction for how we should use it). Addictive desires are surely desires. What is plausible in these writers' accounts is the thought that desire might be a cluster concept, and that there might be dissociable elements within the cluster. For a nice discussion of how this can lead to difficulties in deciding which desire is strongest, see Lloyd Humberstone, 'Wanting, Getting, Having' *Philosophical Papers* 19 (1990), 99–118.

²⁵ The findings here concern inability to delay gratification rather than addiction. See R. Karniol and D. Miller 'Why not wait? A cognitive model of self-imposed delay termination' *Journal of Personality and Social Psychology* 45 (1983) pp. 935–42. It is well established that, after making a choice, subjects tend to view the option that they have chosen as more markedly superior to the alternatives than they did before: a phenomenon that is known as 'post-decisional spreading of alternatives. See J. Brehm and H. Cohen *Explorations in Cognitive Dissonance* (New York: Wiley, 1962).

²⁶ Baumeister et al, 'Free will in consumer behavior: Self-control, ego depletion and choice' *Journal of Consumer Psychology* 18 (2008) 4–13, at p. 9.

²⁷ For good discussion of the features of craving see George Loewenstein, 'A Visceral Account of Addiction' in J. Elster and O-J Skog (eds.) *Getting Hooked*, (Cambridge: Cambridge University Press, 1999) 235–64

How then does resistance work? Our best account comes from studies of delayed gratification. Mischel found that children can resist temptation better if they can avoid reconsideration of their intention to hold out; once they start to reconsider they are likely to succumb.²⁸ Similarly, Gollwitzer has found that forming implementation intentions—that is, intentions that are detailed enough that one hardly needs to think in order to implement them—avoids ego depletion.²⁹

So this gives us some take on the kind of effort required to resist temptation, and on why it is so hard. One is confronted with insistent desires that capture the attention. Since one cannot undermine them, one can only resist them by endeavouring not to reconsider one's resolutions: a form of mental control.

Let me say a little about what this control may be like. It might seem that the best strategy is not to think about one's resolutions at all: form them and then act on them, without so much as contemplating them or the possibility of acting otherwise. Perhaps this is the idea that we have the very strong-willed individual who, as we might say, is never really tempted by the alternatives. However, in many cases it would not work at the very point at which it is needed.³⁰

This kind of unthinking pattern best describes those actions that are automatic. Force yourself to get up at six every morning to go for a run, and after a while it will probably become automatic. The alarm clock will go, you will get out of bed, put on your running kit, and get outside without really giving thought to what you are doing. But the point at which an action becomes automatic is really the point at which will-power is no longer needed. There is good reason for this. At least to begin with, a resolution is typically a resolution to reform one's behaviour into paths which are not automatic. Indeed standardly the automatic behaviour is frequently exactly the behaviour that one has resolved to stop—lighting up a cigarette for instance. If one is to be successful in resisting having a cigarette, and if cigarettes are around, one must constantly monitor whether or not one has picked one up; and one can hardly do that without thinking about cigarettes, and the possibility of smoking them. Successful resolutions cannot work unthinkingly.³¹

So to maintain a resolution like giving up smoking we need something in between full-blown reconsideration and unthinking action. Most resolutions are, I suspect, like

²⁸ For a summary of a large body of work see Mischel 'From Good Intentions to Willpower' in P. Gollwitzer and J. Bargh (eds.) *The Psychology of Action* (New York: The Guildford Press 1996) pp. 197–218.

²⁹ P. Gollwitzer, K. Fujita, & G. Oettingen, 'Planning and the implementation of goals', in R. Baumeister & K. Vohs (eds.) *Handbook of Self-Regulation* (New York: Guilford Press, 2004) pp. 211–228; P. Gollwitzer, U. Bayer, & K. McCullough, 'The Control of the Unwanted', in R. Hassin, J. Uleman, & J. Bargh (eds.), *The New Unconscious* (Oxford: Oxford University Press, 2005) pp. 485–515. This might help explain why people are more likely to be successful giving up alcohol than losing weight. One can form a simple implementation intention never to drink alcohol again, one that will need no further consideration. In contrast one cannot simply decide not to eat again: one will need to be constantly considering what is a permissible amount. For this explanation see G. Ainslie, *Breakdown of Will* (Cambridge: Cambridge University Press 2001) p. 147.

³⁰ Here and in the seven paragraphs that follow I draw on my 'How is Strength of Will Possible?' in S. Stroud and C. Tappolet (eds.) *Weakness of Will and Practical Irrationality* (Oxford: Clarendon Press, 2003) pp. 39–67, which gives a more detailed account; see also my 'Rational Resolve', *Philosophical Review* 113 (2004) pp. 507–35.

³¹ The need for self-monitoring is central to Carver and Scheier's influential feedback account. See C. Carver and M. Scheier, *On the Self-Regulation of Behaviour* (Cambridge: Cambridge University Press 1998).

this. What we need is a state that involves awareness of the resolution, and perhaps of the considerations for which it is held, but which doesn't involve reconsideration. The crucial factor here is that the resolution is not suspended. To remind oneself of one's resolutions is not, by itself, to bring them into question. It is important that it is not suspended. For, as we have seen, once a resolution is suspended, it will all too easily be revised. We thus need a state of awareness that falls short of suspension: what I shall call *rehearsal*.

I speak as though the contrast between reconsideration and rehearsal is a sharp one. In fact, of course, there will be many states in between: what I have marked out are the extremes of a continuum. Moreover, very often mere rehearsal will lead one into reconsideration. This is unsurprising when one's rehearsal leads one to dwell on the benefits to be gained by yielding to temptation; but empirical work shows that the same effect will often come even when one's focus is on the benefits to be gained by holding out.

Can we resist the slide from rehearsal to reconsideration by dint of mental effort? It might seem that this would require an ability to repress thought. The difficulty with such advice is that it is very hard to control one's thoughts directly. Indeed, the effort is typically counterproductive: attempting to repress a thought leads one to dwell on it all the more.³² But need it be that mental control involves such direct repression?

In seeing the possibilities it is useful to look to the advice given by those professionally concerned with the business of resisting temptation. Here is a representative passage from Ignatius of Loyola, founder of the Jesuits:

There are two ways of gaining merit when an evil thought comes from outside: the first ... I resist it promptly and it is overcome; the second I resist it, it recurs again and again and I keep on resisting until the thought goes away defeated ... One sins venially when the same thought of committing a mortal sin comes and one gives ear to it, dwelling on it a little or taking some sensual enjoyment from it, or when there is some negligence in rejecting this thought.³³

Quite what does 'resisting' a thought amount to? It does not seem that Ignatius is calling for out-right thought suppression. Rather he talks of the risks of dwelling on a thought, or of taking some sensual enjoyment from it. The idea seems to be, not that we can keep certain thoughts out entirely, but that we can avoid focusing on them and developing them. Here it does seem far more plausible that we have some control.

I know of no studies on this, but some light might be shed by considering some parallel cases, even if the parallel is far from perfect. Suppose I ask you not to think of the number two. That is almost impossible, and the very effort of monitoring what you are doing makes your failure all the more certain. But suppose I ask you not to multiply by . Unless you are extraordinarily good at mental arithmetic, so that the answer simply jumps out at you, you won't find my request hard to comply with at all. Nor will your monitoring of what you are doing undermine your compliance. Similarly, suppose I ask you not to think through, in detail, the route that you take from home to work. You

³² D. Wegner, *White Bears and Other Unwanted Thoughts* (New York: Viking Press, 1989); J. Uleman and J. Bargh (eds.) *Unintended Thought* (New York: the Guildford Press 1989). Again, it has been suggested that this is connected with the idea of self-monitoring: in order to be sure that one is not thinking about something one needs to monitor that one is not: D. Wegner 'Ironic Processes of Mental Control' *Psychological Review* 101 (1994) 34-52.

³³ Spiritual Exercises, ¶¶ 33-5, in *Personal Writings* (Harmondsworth: Penguin,) p. .

might not be able to resist imagining the starting point; but I suspect, unless you live very close to work, that you will be able to stop yourself somewhere down the track. The point seems to be that there are quite a few steps needed to perform a long multiplication or to imaginatively trace one's route, steps that have to be taken in a particular order, and one is able to exercise some control over such a process.

I suggest that things are typically similar with the thoughts involved in the revision of resolutions. It might be impossible to control whether we entertain the thought of having a cigarette. But it might be possible to control whether or not we go through the procedure that is involved in revising one's resolution not to. This also seems to be the kind of thing that Ignatius has in mind. The sin does not consist in having the evil thought that 'comes from outside'; Ignatius accepted that that is inevitable. The sin comes with what I do with it. And it is here that resistance is hard.

MUDDLING FATALISM AND DETERMINISM

How does this help explain the experience of freedom? As with making a choice, the idea is that one is actively doing something when one is sticking with a resolution. Once again one's prior psychological states—desires, beliefs, intentions, etc.—do not suffice for determining what one will do since they do not determine how much effort one will put in. How does this explain a tendency to embrace incompatibilism? Again one possibility is that people misrepresent this local non-determinism as global. Alternatively, and this is the interpretation I favour, it could be that they tend to misunderstand what determinism would be like, as something that is incompatible with effecting change by employing effort. Let us consider this possibility in more detail.

Fatalism is a common theme in fiction. An excellent, and much quoted, example comes in the speech that Somerset Maugham gives to Death:

There was a merchant in Baghdad who sent his servant to market to buy provisions and in a little while the servant came back, white and trembling, and said, Master, just now when I was in the market-place, I was jostled by a woman in the crowd and when I turned I saw it was Death that jostled me. She looked at me and made a threatening gesture; now, lend me your horse, and I will ride away from this city and avoid my fate. I will go to Samarra and there Death will not find me. The merchant lent him the horse and the servant mounted it, and he dug his spurs in its flanks and as fast as the horse could gallop he went. Then the merchant went down to the market-place and he saw me standing in the crowd and he came to me and said, Why did you make a threatening gesture to my servant when you saw him this morning? That was not a threatening gesture, I said, it was only a start of surprise. I was astonished to see him in Baghdad, for I had an appointment with him tonight in Samarra.³⁴

³⁴ S. Maugham *Sheppey: A Play in Three Acts* (London: W. Heinemann, 1933), Act III. It is quoted by many authors from John O'Hara on. The origin of the story appears to be from the 9th Century author Fudail ibn Ayad; see Idreis Shah, *Tales of the Dervishes* (London: Jonathan Cape, 1967) p. 191. I am indebted here to J.H. Sobel 'Notes on "Death Speaks"' at <http://www.scar.utoronto.ca/~sobel/PuzzlesDEATHSPEAKS.pdf>

The servant's effort and enterprise will make a difference to the immediate outcome: he will get on the horse and ride to Samarra. But whatever he does he won't be able to escape the ultimate end to which fate has assigned him. The theme is a familiar one in classical mythology, and more recently in those time travel stories in which an agent is involved in a project in which we know they must fail—trying to kill their own younger self for example. In terms of Seligman's continuum, one's fate belongs at the extreme end of what one can't change, even beyond the fruitless attempt to lose weight.

I suspect that a common idea is that if determinism were true, everything would be like that: effort would make no difference to outcome. On this view determinism would be a form of global fatalism, in the sense that one would be unable to change one's actions. Of course, there is a difference between the classic fatalist cases and determinism. As I mentioned, classic fatalism involves frustration at a distance. One performs one's basic action, only to find that it doesn't help: the servant gets to Samarra, only to find that that death is there waiting for him. If determinism is a universal form of fatalism, it will also bring a more immediate frustration: if one has been determined to do otherwise, then one will not even perform the basic action that one is aiming to perform, no matter how much effort one puts in. But this just makes the fatalism all the more pressing, since it begins at home. No exercise of effort will change anything. Every action is crowded at one end of Seligman's spectrum.

I hope it is clear how this involves a mechanistic picture. It is certainly one in which it is reasonable to think of us as pawns in the hands of impersonal forces. I hope that it is equally clear that as an interpretation of determinism it is simply a mistake. It is reminiscent of the Lazy argument used against the Stoics: if one is fated either to recover or to die, what use is there in calling the doctor? Chryssipus's response, whether or not it answers all the worries, is apposite to our concerns here: to say that it is determined that Oedipus will be born to Laius is not to say that this will be so 'whether or not Laius has intercourse with a woman'.³⁵ Having intercourse is a necessary condition for having a son. If it is determined that Oedipus will be born to Laius, then it is determined that Laius will have intercourse; the two are 'co-fated', to use Chryssipus's term. But that doesn't undermine the causal efficacy of the act.³⁶

Similarly if we got conclusive proof of the truth of determinism, we wouldn't throw Seligman's book away. Seligman is talking about how much effort is needed to effect different kinds of change. In a deterministic model it is still true that different changes require people to put in different amounts of effort, and, crucially, that the effort they put in makes a difference. If there hadn't been the effort, the outcome would have been different. So it is still true that in that sense people act: their efforts are not thwarted by forces that ensure an inevitable outcome.

But despite being a mistake, my suggestion is that this is the view that reading much deterministic literature tends to induce in people. They understand the impact of determinism in terms of a distinction—Seligman's distinction between what you can change and what you can't—that they already have. And it leads them to think that every

³⁵ As quoted by Cicero, *On Fate* 30, translation in A. Long and D. Sedley, *The Hellenistic Philosophers* (Cambridge: Cambridge University Press, 1987) p.339.

³⁶ For discussion of Chryssipus's argument see Susanne Bobzien, *Determinism and Freedom in Stoic Philosophy* (Oxford: Clarendon Press, 1998) Ch. 5. As she points out (pp. 231–3), Chryssipus is not really concerned to answer the question of how something that is determined can still be up to us: what I earlier called the more global skepticism.

attempt to influence things by putting in effort falls on one side of that distinction. To use the psychological terminology, it undermines their self-efficacy.³⁷

DETERMINISM AND MORAL MOTIVATION AGAIN

My contention then is that reading determinism passages will tend to increase subjects' scepticism about whether their efforts will have any impact. In particular, it will lead them to think that struggling against their inclinations will be futile. This is not to say that they will cease acting, but that they will think of their actions as themselves determined by their beliefs and desires.

Inducing belief about the futility of struggle is a classic seduction strategy. Witness how Rodolphe overcomes any remaining resistance from Emma Bovary by describing how fate works to draw lovers together:

Come what may, sooner or later, in six months, ten years, they will be together, will be lovers, because Fate ordains it, because they were born for each other.

And now applying the lesson to them:

Look at us ... why did we meet? By what decree of Fate? It must be because, across the void, like two rivers irresistibly converging, our unique inclinations have been pushing us towards one another.³⁸

Now we have clearly reached the moral sphere. Where people would normally think that they could resist the inclination to cheat, or the inclination to prank with hot salsa, a belief in fate will make them more sceptical. And this will lead them not to try, or to give up more readily if they do.

Here then is how I suggest we interpret the experiments that show that a belief in determinism is morally demotivating. The effect of such a belief will be relatively direct. Subjects do not need to think about the nature of morality, and of whether the existence moral responsibility requires the falsity of determinism. Indeed the considerations that move them are not directly about morality at all. Rather they are just about whether putting in effort will make any difference. Morality is just one instance of the more general phenomenon of resisting contrary inclinations. And that provides a way of evaluating whether my conjectures are true. For if they are, the effect should be just as marked on cases that do not involve morality at all. Dieters who read deterministic passages should be more ready to break their diets.

If this is right, then there is something very interesting about subjects' responses once they take fatalism on board. It is not as though they stop acting altogether: they cheat; they give salsa to those who loathe it. Rather, they stop exercising self-control.

³⁷ In general *self-efficacy*—one's belief in one's own degree of control—is extremely important in explaining behaviour. For a general overview see A. Bandura, 'Exercise of Personal Agency Through the Self-Efficacy Mechanism' in R. Schwarzer (ed.) *Self-Efficacy* (Bristol PA: Taylor and Francis, 1992) pp. 3–38.

³⁸ G. Flaubert, *Madame Bovary*, trans. G. Wall (Harmondsworth: Penguin, 1992) Part 2, Chap. 8. Rodolphe's approach is subtler still, since he suggests that it is merely the pressure of conventional morality that will be overcome by fate; he doesn't acknowledge to Emma that she is struggling against her conscience.

People see themselves as working on two levels. They are moved by certain prior beliefs, desires and intentions; but they have an ability to control whether they will actually act on those states.³⁹ If a belief in the non-existence of free will, and the truth of determinism, impacts primarily on self-control, then this suggests that it is self-control that provides a primary source of the experience of freedom.

More reassuring for philosophers is that these concerns at least leave us with no need to propogate a Noble Lie of indeterminism. What is needed is a more careful—a more *truthful*—account of what determinism involves.

Dept of Linguistics and Philosophy
MIT

holton@mit.edu

³⁹ This two level view is a pervasive feature of our psychological and moral thought. It comes up, for instance, in the legal doctrine of provocation. For discussion see Richard Holton and Stephen Shute, 'Self-Control in the Modern Provocation Defence', *The Oxford Journal of Legal Studies* 27 pp. 49–73.