



Addictive actions

Edmund Henden

To cite this article: Edmund Henden (2013) Addictive actions, *Philosophical Psychology*, 26:3, 362-382, DOI: [10.1080/09515089.2012.660688](https://doi.org/10.1080/09515089.2012.660688)

To link to this article: <https://doi.org/10.1080/09515089.2012.660688>



Published online: 10 Apr 2012.



Submit your article to this journal [↗](#)



Article views: 430



View related articles [↗](#)



Citing articles: 1 View citing articles [↗](#)

Addictive actions

Edmund Henden

It is common to think of addiction as involving behavior which in some sense is “out of control.” But does this mean addictive actions occur because of compulsion or because of ordinary weakness of will? Many philosophers argue that addictive actions occur because of weakness of will, since there is plenty of evidence suggesting that they are not caused by irresistible desires. In fact, addicts seem, in general, to perform these actions freely in the sense of having the ability to refrain from doing so. In this paper I argue, first, that it is not the addiction-as-compulsion view that is mistaken, but rather the view that irresistible desires are a defining feature of compulsion. Second, drawing on some results in addiction neuroscience, I construct and defend a new analysis of compulsivity that distinguishes addictive from weak-willed actions in a way that is consistent with the view that addictive actions are performed freely.

Keywords: Addiction; Compulsion; Free Action; Irresistible Desire; Weakness of Will

1. Introduction

What do we mean when we explain a certain type of behavior by saying that it occurs because of “addiction”? It is common to think of addiction as involving behavior which in some sense is “out of control.” But does this mean addictive actions occur because of compulsion or because of ordinary weakness of will? It has been a standard assumption in the philosophical literature that an action is compulsive only if it is caused by an irresistible desire (Watson, 2004a). However, as is frequently pointed out, there is plenty of evidence for saying that addictive actions are not caused by irresistible desires. For example, addicts often quit drugs without assistance and addiction does not eliminate personal responsibility (Benn, 2007; Levy, 2006; Wallace, 2006). In fact, addicts seem, in general, to perform their addictive actions freely in the sense of having the ability to refrain from doing so. Despite the popularity of the addiction-as-compulsion view in the biomedical sciences, many philosophers therefore conclude that addiction must be a sub-variety of weakness of will rather than compulsion, since weak-willed actions are generally believed to be

Edmund Henden is research fellow in philosophy at the Centre for the Study of Mind in Nature (CSMN), University of Oslo, Norway.

Correspondence to: Edmund Henden, CSMN, IFIKK, Postboks 1020, Blindern 0315 Oslo, Norway.
Email: edmund.henden@ifikk.uio.no

performed freely. The trouble with this inference, however, is that it might not be the addiction-as-compulsion view that is mistaken, but rather the view that irresistible desires are a defining feature of compulsion. In fact, this is precisely what I shall argue in this article. Drawing on some results in addiction neuroscience, I shall construct and defend a different analysis of compulsivity that distinguishes addictive from weak-willed actions in a way that is consistent with the view that addictive actions are performed freely.

In the first part of this article, I shall provide evidence that addictive actions do not, in general, occur because of irresistible desires. However, unlike those who take this to show that addictive actions must be weak-willed rather than compulsive, I shall argue that addictive actions do not occur because of weakness of will either. In the rest of the article I shall develop an account of compulsive action according to which an action is compulsive, not because it is caused by an irresistible desire, but because it is part of a habit the compulsive person would find extremely difficult to discontinue even if she made a sincere effort to do so. The reason she would find it difficult is because the habit's continuation counterfactually depends on a specific type of dissociation in her motivational system which disrupts the normal psychological functioning of her deliberative and volitional processes. That notwithstanding, I shall argue that it is still within her powers to refrain from the compulsive actions that form parts of this habit. Indeed, on this account, there is plenty of evidence for saying that addictive actions are compulsive. In the final section of the paper I defend this account against three objections.

Before I begin, let me clarify some basic terminology. To keep things simple, the term 'addictive action' will be taken to mean a drug-seeking or drug-taking action that can be explained in terms of the agent's addiction to a particular drug. The notion of "acting freely" is, of course, notoriously difficult. To avoid getting entangled in the free will controversy, I shall simply assume that saying that an agent is free with respect to a particular action at some time means that she has the ability to refrain from that action at that time. I am not claiming that possessing such ability is sufficient for free will.

2. Addiction, Compulsion and Weakness of the Will

As mentioned above, a standard assumption in the philosophical literature is that an action is compulsive only if it is caused by a desire that is irresistible (Watson, 2004a). A "desire" is usually understood as a motivating state that involves an attitude about some content, which is capable of rationalizing actions. By this is meant that, if the object of a desire is the state of affairs that one acts in a certain way, one is consciously aware of, in a broad sense, some favorable representation of that state of affairs (or at least that there are no systematic barriers to such a representation becoming an object of one's conscious awareness) which is capable of moving one to try and bring the world into conformity with that representation.¹ In what follows, I shall assume this account is correct. Given this account, *the addiction-as-compulsion*

view holds that addictive actions are compulsive because they are caused by irresistible desires to take drugs (Frankfurt, 1988). The view that a desire is “irresistible” at some time t for an agent S means that S is unable at t to resist acting on that desire.

The modal describing this inability should be relativized to the circumstances that obtain at the time at which the desire is irresistible to allow for the possibility that although a person might lack this ability to resist at particular times, or most of the time, she may still have it under some exceptional circumstances (Mele, 1990; Zaragoza, 2006). For example, a compulsive hand washer who is unable most of the time to resist acting on a desire to wash her hands might still be able to resist acting on that desire if someone threatened to kill her child if she did. Now, being unable to resist acting upon a drug-oriented desire means that one has lost one’s capacity for rational self-control. According to a standard philosophical view, rational self-control should be understood in terms of reasons-responsiveness, where a loss of this capacity implies a lack of susceptibility to countervailing reasons (Fischer & Ravizza, 1998; Watson, 2004b). This means that an addict can be said to have lost this capacity at some time t if a certain sort of counterfactual is true of her: if she were presented with what she took to be good and sufficient reasons for not performing an addictive action at t , she would still perform that action at t .

There appear to be two main sources of evidence for the view that addicts satisfy this counterfactual. First, there is the observation that they often seem to seek and use drugs in spite of judging that they have better reasons to refrain. That is, they bring major harms down on themselves, including emotional distress, legal and financial problems, health problems, etc., and often report feeling miserable, that their lives are in ruins, and that they regret ever having started down the road that has led them to this state (Robinson & Berridge, 1993). Second, there is the observation, generally acknowledged by both researchers and practitioners in the field, that addicts in many cases consciously and sincerely try to resist the tendency to perform the addictive actions (often with great costs to themselves), that they experience this as difficult, and that their attempts frequently fail, giving rise to reports that their actions “feel out of their control” (Heather, Tebbutt, Mattick, & Zamir, 1993; West, 2006). Given these observations, one might argue, inference to the best explanation suggests that addicts must have lost their capacity for rational self-control. However, in general, when we explain a person’s failure to successfully perform a certain action because she finds its performance hard, we do not necessarily want to imply that she failed because it was *impossible* for her to perform the action successfully. Often, there will be reasons other than loss of capacity for why she finds it hard. For example, she may have been distracted or exhausted, or have simply misjudged the effort required to perform the action successfully.

Could something similar be true of addicts’ capacity for rational self-control? What seems clear is that the cited evidence does not rule out this possibility. When it is added that there appears to be evidence that many addicts do in fact successfully exercise rational self-control, the view that irresistibility is a defining feature of addiction seems to have little support. Thus, several studies show that financial

concerns, fear of arrest, values regarding parenthood and many other factors which influence decisions in general, often bring addicts' drug-oriented behavior to a halt (Heyman, 2009). An example given by Heyman (2009) serves to illustrate this point. Heyman describes an addiction treatment program combining counseling sessions aimed at assisting cocaine addicts on practical matters (such as how to find a better job, improve relations with family members, and obtain better training) with an exchange system offering vouchers for metabolic evidence of drug abstinence. The vouchers, which the addicts earned if their urine tests were drug-free, could be traded in for desirable but modest goods, such as movie passes and sports equipment. For each additional week of abstinence, the value of the vouchers would be increased. Conversely, if the addicts tested positive for drugs, the value would be set back to its initial, lowest level. Compared with two control groups where some received traditional psychological counseling and others received vouchers independent of whether they had been abstinent, the contingency group showed significantly higher abstinence rates.² A subsequent study found that at every follow-up date after the removal of the contingency, the percentage of drug-free voucher subjects rose from 60 to almost 80 percent over the year (Heyman, 2009, p. 107). Clearly, this is not the result one would expect if addiction implied a loss of the capacity for rational self-control. What it suggests is that many addicts *can* be persuaded to exercise this capacity successfully if the reasons they are given appear to them to be good and sufficient.

If addictive actions do not occur because of irresistible desires for drugs, should we conclude that they occur because of *weakness of will* instead? I shall later suggest an argument why we should not, but at this stage I just want to mention two reasons for remaining somewhat skeptical toward this conclusion. First, the phenomenology of addiction seems different from the phenomenology of weakness of will.³ As an illustration, consider Austin's (1979) famous example of a case of weakness of will. Austin describes a situation at High Table where he helps himself to two segments of ice cream in spite of a belief that it would be better to abstain. He remarks: "but do I lose control of myself? Do I raven, do I snatch the morsels from the dish and wolf them down, impervious to the consternation of my colleagues? Not a bit of it. We often succumb to temptation with calm and even finesse" (1979, p. 198). What is philosophically puzzling here is that Austin appears to voluntarily succumb to the temptation to have more ice cream in spite of his judgment that, all things considered, it would be better to abstain. By "voluntarily succumbing," I mean his *allowing himself to give in* to his desire for more ice cream by intentionally not making any sincere attempt to resist it. That is why we think his will is weak.⁴ What makes his action philosophically puzzling is that we seem to have no reason to expect it *not* to reflect his judgment that it would be better to abstain. That is, what is puzzling is that he seems to be cooperating in his own failure to resist acting on his desire for more ice cream—a desire he judges he has most reason to resist and could have resisted if he had sincerely tried—by not deciding to sincerely try. By refraining from making that decision he can indulge his desire, let himself be moved by it, in spite of his judgment that abstaining would be better. However, the phenomenology of addiction seems very different. Addicts frequently report experiences of failures to

resist acting on their drug-oriented motivation in spite of *trying hard* to resist (Heather et al., 1993). That is, unlike Austin, not only do they often decide to try to resist their drug-oriented motivation; they actually sincerely try to resist it. Yet, they frequently fail, something that gives rise to feelings of compulsion (West, 2006, p. 75). This suggests that their failures depend, at least in part, on factors beyond their conscious, intentional control. Phenomenologically, it does not seem correct, therefore, to hold that addicts generally give in to their drug-oriented motivations because their wills are weak.

Second, one central feature of weakness of will is that weak agents are believed to be rationally criticizable for succumbing to temptation. That is, we believe it reasonable to expect (and demand) that they resist acting on their desires. It has been common in the philosophical literature to assume that this implies that we ascribe to weak agents capacities for rational self-control (Davidson, 1980; Smith, 2004).⁵ The reason weak agents are believed to be rationally criticizable, then, is because we hold them responsible for not properly exercising capacities we deem them to have. Suppose the autonomy impairment characteristic of addiction can be explained in terms of weakness of will. Then, other things being equal, there should be little difference with respect to the criticizability of addicts' and weak-willed *non-addicts'* failures to control their drug-oriented actions. However, this implication does not seem supported by intuition. Compare a heroin addict and a weak-willed non-addict "experimenter" who both inject themselves with heroin. Suppose that both weigh and assess their available relevant reasons and conclude that it would be best to abstain. In other words, both act against their best judgments. We may assume that both possess a capacity for self-control, and that their drug-oriented motivation is resistible at the time of action. They could plausibly be criticizable for their failure to control themselves, but are they *equally* criticizable? Clearly, they are not. We are more inclined to criticize the non-addict experimenter than the heroin addict. The reason seems obvious: while we assume it would be relatively easy for the former to have refrained from injecting herself with heroin ("she could just have tried a bit more!"), we assume it would be comparatively much harder for the latter to refrain from doing the same. This difference affects our attitudes to the two cases, suggesting that our attitudes to addiction importantly differ from our attitudes to weakness of will.⁶

I am not claiming that the considerations above provide conclusive evidence against the view that addiction is a species of weakness of will. Yet, they do make this view appear intuitively less attractive. In the rest of this paper I shall argue that, although addiction does not occur because of irresistible desires for drugs, it should still be considered a species of compulsion.

3. Desires, Habits, and Inclinations

In most clinical descriptions of compulsive behavior, what tend to be emphasized are the following observable features. First, compulsive behavior is strongly

cue-dependent in the sense that it is regularly triggered by certain situations, places, or people associated with the type of behavior in question. Second, compulsive persons feel repetitively driven to perform it, often in spite of themselves. That is, reports of feelings of compulsion are common. Third, if compulsive persons sincerely try to refrain from acting upon their compulsive motivation it becomes, they report, increasingly difficult over time (Abramowitz & Deacon, 2005; de Silva, 2006). It seems fair to say that none of these features are essential to weak-willed actions. That is, unlike compulsive actions, these actions tend neither to be strongly cue-dependent, associated with sincere attempts to refrain, nor with feelings of compulsion. Since, however, there appears to be plenty of evidence that *addictive* actions tend to have these features, a plausible hypothesis might be that addiction has more in common with compulsion than with weakness of will.⁷ It could be objected, perhaps, that the one thing that distinguishes addictive from compulsive actions is that while the former are flexible and adaptive, the latter are usually carried out according to rigid rules or in a stereotyped fashion. However, this difference seems inessential. The fact that an action is *not* carried out according to rigid rules or in a stereotyped fashion does not, in general, imply that it cannot be compulsive. For example, kleptomania is standardly understood as involving compulsivity even if the actions of kleptomaniacs can exhibit substantial degrees of flexibility and adaptiveness.

Although there appears to be wide agreement on many of the observable features of compulsive behavior in the clinical-scientific literature, it is difficult to find any precise metaphysical account of the notion of compulsivity itself. Yet, in order to decide whether or not addictive actions are compulsive, such a metaphysical account might be helpful. In philosophy, metaphysical discussions of compulsive action often start by identifying some special feature of the agent's desire to perform this action—a feature that is assumed to be essential for, or constitutive of, compulsivity—and then proceeds to analyze compulsive actions in terms of this feature. “Irresistibility” is, of course, the prime example of such a feature, presumably because many philosophers see compulsive action simply as the negation of free action. A general difficulty for this kind of approach is that it is often very difficult to determine, based on observational evidence, whether the relevant feature is actually present in real, clinical cases of compulsive action (see last section). I want to start instead by focusing on an observable feature of compulsive actions (often left out in philosophical discussions), namely the regularity with which they occur. That is, the clinical-scientific notion of compulsivity seems to refer, not simply to isolated events, but to *patterns of behavior* performed on a regular basis, in characteristic circumstances, which the agent finds it difficult to override by intentional effort.

Now, focusing on patterns of behavior instead of isolated instances does not, of course, rule out explanations in terms of features of the agent's desires. One kind of proposal might be that what explains such patterns (as well as the difficulties of breaking them) is that these desires persistently and constantly direct the agent's attention to their objects in terms of some evaluative category, such as, say, “pleasant” or as “something to be enjoyed,” independently of practical deliberation

and voluntary control (Wallace, 2006, p. 180; Watson, 2004b, p. 72). By dominating her conscious experience and breaking her concentration on other things, they succeed in redirecting her will, or “seducing” her to repetitively perform actions which thereby take on a compulsive character. However, it does not follow that because it is hard to resist *having* certain desires (since they come to one unbidden and spontaneously), that it is equally hard to *refrain from acting* upon those desires. Desires work through the agent’s intentional system (Morse, 2000). An action is not, therefore, simply a causal effect of a desire. In fact, the occurrence of inappropriate desires, e.g., of a violent or sexual nature, appears to be quite common in the general population (de Silva, 2003, p. 32). While most normal persons with an ordinary capacity for rational self-control successfully refrain from acting upon such desires if they sincerely try to do so (i.e., if they fail to refrain, it is more likely to be due to weakness of will than to compulsion), compulsive persons regularly fail to refrain from compulsive actions in spite of sincerely trying to refrain. Is that because their desires *are always more forcefully and intensely* directing attention onto the pleasurable features of their objects than those of normal persons?

I want to explore another possibility. Not all behavior needs to be related to desires in order to count as intentional. Consider habitual behavior, for example, such as my habit of taking a particular route to work every day. It seems plausible that taking this route is something I do intentionally; yet what explains this behavior is not any prior desire, deliberation, or decision. What explains it is just habit. But what kind of explanation is that? It does not seem to resemble an ordinary causal explanation, since in specifying the explanans (the habit) we specify the explanandum itself (the action which, amongst other actions, is to be explained). So the habit and the action it is meant to explain are not logically distinct entities. One possibility is that habit explanations are a kind of constitutive explanation. That is, we explain the action by saying it is *part of*, or *constitutes*, something else, namely a certain behavior pattern which is regularly performed in characteristic circumstances (Pollard, 2006). This should not be taken to imply that habitual behavior does not involve intentions. According to one plausible view, intentions are executive attitudes toward goal representations (or, as some would put it, “plans”), the activation of which is necessary for initiating, sustaining, and guiding behavior; that is, for constituting behavior *as* intentional (Mele, 2009). It follows that since taking this particular route was something I did intentionally, I must have had an intention that explains why doing it was intentional under a certain description, i.e., some goal representation that had as part of its content this act under that description. However, in contrast with ordinary intentions which I intentionally form or decide to form as the result of prior conscious states or processes (e.g., beliefs, desires, deliberations), the goal representation guiding an instance of habitual behavior can be activated non-consciously by environmental cues as the result of conditioning, independent of prior conscious states or processes. Habitual behavior, then, provides a counterexample to the claim that behavior must be appropriately related to desires in order to count as intentional.

Now, habitual and compulsive behavior may seem similar in at least this respect: like the notion of a habitual action, the notion of a compulsive action appears to refer

to a behavior pattern that has that action as a part, which is regularly performed in characteristic circumstances. It therefore seems plausible that *explaining* an action in terms of compulsion would also seem to require seeing such action as part of such a behavior pattern. But is there any evidence that compulsive actions occur in the absence of desires? In fact, there is. Consider first obsessive-compulsive disorder, which is one of the most notable compulsive disorders. Many clinical researchers believe that the repetitive behavior patterns of subjects suffering from this disorder serve the function of regulating affect or reducing negative feelings. Yet there is no logical connection between the description of these subjects' actions and the goal they are intended to achieve. In effect, the subjects themselves often have no idea why they repeatedly perform actions such as washing or checking (Abramovitz & Deacon, 2005; de Silva, 2006). That is, not only do they fail to see anything of value in the objects of their compulsions, their own actions seem to them excessive, unpleasant, and pointless. Assuming that desires reveal the light in which actions are or appear favorable from the agent's own perspective, it would be incorrect to ascribe to these agents a *desire* to act in this way.

In fact, perhaps more surprisingly, desires also seem less important in understanding *addiction*. In a series of influential papers, the neuroscientists Kent Berridge and Terry E. Robinson (hereafter B&R) argue that, although addicts are often motivated to obtain pleasure or relieve withdrawal, a vast amount of compulsive drug-seeking and drug-taking behavior remains to be explained. Thus, there is plenty of evidence that addicts often continue to seek and take drugs even when no pleasure can be obtained, and even in the absence of withdrawal (Robinson & Berridge, 1993).⁸ B&R take this to pose a problem for the traditional view in psychology according to which the affective-cognitive and action-driving components of motivation are mediated by a single psychological process and neural substrate. On this view, incentive value (the degree to which a goal or stimulus is action-driving) is explained in terms of the affective-cognitive component of motivation, i.e., in terms of how much subjective pleasure, goodness, utility, or other positive value the incentive is expected to bring. B&R call this affective-cognitive component "liking," (Robinson & Berridge, 1998) and provide evidence that the psychological process and neural substrate responsible for determining liking are separable from the psychological process and neural substrate responsible for determining incentive value, or what they call "wanting."⁹ While normally the affective-cognitive and action-driving components of motivation go together, so that we "want" the things we "like" (e.g., the pleasure associated with some environmental cue or circumstance serves as a trigger to activate and direct wanting), in addiction they come apart, making addicts want things they do not like. The reason, according to B&R's theory of incentive-sensitization, is because repeated drug use "sensitizes" the brain regions involved in the process of wanting—making them more easily activated by drugs or cues reminding the addict of consumption—but not the brain regions involved in the process of liking.

Although the scope of the concept of liking, as B&R use it, is wider than that of the philosophical concept of desire (e.g., it includes preconscious affective-cognitive

processes in humans as well as animals), “desiring” something (in the philosophical sense) does seem to imply liking it (in B&R’s sense). Thus, an agent cannot fail to have a liking attitude (in B&R’s sense) toward a content she represents as counting in favor of, e.g., as desirable, pleasurable, enjoyable, or good. If it is added that agents sometimes want things they *do not* like, it follows that wanting something (in B&R’s sense) does not imply desiring it (in the philosophical sense). In other words, “wantings” can occur in the absence of any corresponding affective-cognitive states representing their contents as counting in favor of, e.g., as desirable, pleasurable, enjoyable, or good. Perhaps one way of distinguishing wanting (in B&R’s sense) from desiring (in the philosophical sense), might be in terms of the notion of “knowledge.” If there is, as some philosophers think, a kind of constitutive relation between desire and the reflective attitude involved in knowing what one desires, we might say that this relation is absent in cases of wanting (Shoemaker, 1996).¹⁰

Now, consistent with these ideas, let us suppose that human motivation consists of two separable components, an affective-cognitive component, call it “desire,” and an action-driving component, call it “inclination.” In contrast to desires, inclinations are believed by many philosophers to be brute dispositions caused by bodily mechanisms, a kind of happening external to the agent that arises independently of principles of reason and justification (Quinn, 1995). An alternative view, however, might be to see them as *the beginnings* of primitive actions the agent has already undertaken before they are manifested in overt behavior (Schapiro, 2009). According to this view, the feeling of inclination is simply the agent’s awareness of her exercising her capacity to *initiate* action in response to some environmental cue that has grabbed her attention. By “primitive action,” I mean that the capacity in question is typically exemplified in less than full blooded agencies. Small children, for example, as well as many nonhuman animals, clearly have basic agential capacities in the sense of causal powers to bring about goal-directed movements of their bodies, capacities which operate on the basis of simple principles of conditioning rather than on the basis of principles of reason and justification. These capacities may well be an aspect of our animal nature that survives into adulthood where it continues to shape our behavior (Steward, 2009). In fact, if inclinations can be thought of in this way, it is interesting to speculate whether they might be a kind of “simple” intentions, similar to the intentions guiding habitual behavior. Thus, just as with the latter, such intentions might be triggered when our primitive agential capacities are being directly engaged in response to environmental cues or circumstances as the result of conditioning, independently of prior desires, deliberation, and, in many cases, even conscious beliefs.

This view is not without evidence. For example, there is research suggesting that “drives” or “impulses” involve parts of the brain associated with the execution of primitive actions rather than with cognitive-motivational states, such as desires (West, 2006, p. 148).¹¹ If inclinations can be understood as simple intentions, that would surely account for the motivational force they have on occasions when the agent’s attention is drawn to them because they clash with her will. Rather than the feeling that she *has a desire* to perform a certain type of action in spite of her will,

the subjective manifestation of such conflicts is likely to be the feeling that she *is about to perform* a certain type of action in spite of her will, that a certain type of action is *underway*—a type of action that is difficult to stop once started. These difficulties might be thought to increase the more entrenched the action is due to repetition and reinforcement, and the more they increase, the more recalcitrant the action might seem to reason. Now, consider first the following analysis of compulsive actions that takes this habitual element into account:

- (Df.1) A particular action *A* is performed compulsively by (an agent) *S* if and only if *A* satisfies the following conditions: (i) *A* is part of a behavior pattern that *S* regularly performs in characteristic circumstances; (ii) *S* has inclinations to perform *As* which drive *S*'s performance of *As*.

Conditions (i) and (ii) say, roughly, that the attribution of a compulsive *A* requires that *S* has an inclination to *A*, but also a certain sort of history with respect to *As*—that compulsivity cannot be attributed on the basis of a single action but only on the basis of patterns of behavior the agent has gone in for with some past regularity. The latter follows since part of what makes actions compulsive is their repetitiveness, the fact of being regularly triggered by certain situations, places, or people. This is the respect in which compulsions are similar to habits: just as it does not make sense to speak of one-off habitual actions, it does not make sense to speak of one-off compulsive actions either. Now, this simple analysis is plainly not sufficient for compulsive actions. For one thing, it does not distinguish compulsions from habits, since habits *are* behavior patterns regularly performed in characteristic circumstances. Yet it is not until one's control of a habit is impaired that it might become compulsive. In addition, it may appear to rule out a perception of addiction as a sub-variety of compulsion, since addictive actions often seem driven by strong *desires* for drugs.¹² In the next section I shall propose an account of what it means to have impaired control over a habit and argue that what makes an action compulsive is that it is part of such a habit.

4. Compulsion as Impaired Habit Control

Let me start by returning to a point made earlier, based on the interpretation of B&R's research, namely that human motivation consists of two separable components: desire and inclination. Following B&R, we may then suppose that while desires and inclinations normally go together (i.e., an agent's desires explain her inclinations by activating or directing them), in some cases they diverge. In those cases the agent's actions can be triggered directly by environmental cues or circumstances via a process that is dissociated from her desires, deliberation, and, in many cases, even conscious beliefs. The suggestion I now want to make is that compulsive patterns of behavior essentially involve such dissociations in the agent's motivational system, which may drive her to repetitively perform actions of a certain type in spite of having no desire to do so. They are constitutive features of these behavior patterns and hence of compulsive action. In fact, dissociative *experiences*, such as feelings of "standing

outside oneself while acting,” are frequently reported across a range of compulsive phenomena, including addictions to drugs (Evans & Coventry, 2006). Consider now the following additional requirement for an action’s being compulsive that incorporates this element.

- (iii) S would have continued to perform As despite (a) having desires to refrain from As and no desires to perform As, and (b) judging that there are better reasons to refrain from As.

Condition (iii) captures in counterfactual terms in what sense compulsivity implies a dependency on the object of compulsion independent of S’s desires and reasons. It implies that the link between action and inclination has a measure of robustness, in the sense that it survives in a whole range of counterfactual circumstances, including circumstances in which S has a desire to refrain from A and no desire to perform A, and judges that there are better reasons to refrain from A. Since the idea of “control over action” itself is a modal notion that also must be robust in this sense, it follows that even if S, in the actual circumstances, has a strong desire to A and believes there are better reasons to A, her performance of A is not *controlled* (in any robust sense) by her desire or reasons, since in possible circumstances in which she were to *lack* them, she would still perform As.¹³ It is worth noting here that this does not rule out that strong desires for drugs may play a role in explaining addiction (assuming the latter is a sub-variety of compulsion). For example, it seems plausible that strong desires explain why individuals start using drugs on a regular basis. However, the importance of strong desires seems, in general, to subside once addiction gets a foothold (Lyvers, 2000; Robinson & Berridge, 2003; West, 2006). When this happens, desires do not appear to be what is primarily driving the behavior anymore.

However, (i)–(iii) are not yet sufficient for compulsivity. One may have inclinations to do various things, even do them in the absence of any desire and in spite of judging that one should not, without that necessarily implying that one does them compulsively. Part of what makes actions compulsive is the likelihood of continuing to perform them despite *sincere attempts to refrain*. That is one thing that distinguishes compulsive from weak-willed actions: compulsive persons repeatedly fail to refrain from compulsive actions because they find it too hard to refrain from them, not because they find their will to refrain is too weak (see section 2). But what is “hard” here supposed to mean? In his classic paper on weakness of will, Watson (2004a) argues for conceiving of “compulsion” as a normative notion, in the sense that it is relative to the capacities of resistance of “the normal person” or “typical adult in our society.” Since he assumes that irresistibility is a defining feature of compulsion, what is normative, on his view, is the notion of irresistibility itself. Actions are compulsive, in other words, if they result from desires which are impossible to resist given the capacities of the normal person.

I claim, for reasons to which I return in section 5, that irresistibility is not a defining feature of compulsion, but Watson is right to see a normative element in the notion of “compulsion.” What introduces this normative element, I want to suggest, is the notion of *trying*. By “trying” in this context, I mean in the ordinary language

sense of “making an effort.”¹⁴ If you fail to do something for which you have an ability, you usually fail because you make an insufficient effort (you’re not really trying hard enough). By contrast, if you fail because you lack the ability, the reason you fail is lack of ability, not insufficient effort.¹⁵ In other words, if compulsive persons generally possess the ability to refrain from the compulsive action, the reason they fail to exercise this ability successfully might be put down to insufficient effort. But what is “sufficient effort”? Normally, we have ways of determining the correct answer to questions such as: “did she try hard enough?”—presumably because there are certain shared expectations and norms guiding what count as sufficient effort in various contexts. For example, someone might make what seems a sufficient effort in a particular context, relative to ordinary standards to perform some action, in the sense of making an effort such that if any normal person had made that effort to perform the same type of action in similar circumstances, it would be reasonable to expect the performance of her action to succeed. Of course, this does not mean *her* effort was sufficient relative to what was actually required of her at that time. Relative to that, her effort might indeed have been insufficient. Perhaps she faced some obstacle particular to her which evades normal persons. The reason she failed to perform the action could still be an insufficient effort on her part rather than lack of ability, since if she had made an even greater effort (which, we may assume, was consistent with her actual ability), she would have performed the action successfully.

Now, I want to suggest that what counts by ordinary standards as sufficient effort to refrain from actions for which one has inclinations or desires of strength *s* to perform, is the level of effort which, other things held constant, would have been sufficient for a normal person to successfully refrain from performing actions of the same type if she had desires and inclinations of strength *s* to perform them. I am assuming here, of course, that “a normal person” is someone with a normal capacity for self-control whose motivational system is congruent and who succeeds in refraining from acting upon her desires and inclinations if she makes a sincere effort to do so. That is, when such a person *fails* to refrain from acting upon her desires and inclinations, it is always due to some form of weakness of will. She is simply not making what counts as a sufficient effort by ordinary standards. We can therefore say of an agent who sincerely and repeatedly makes such efforts, but who fails—again and again—to staunch her performance of a certain type of action, that she exhibits a compulsive behavior pattern. Moreover, the fact that her efforts count as sufficient by ordinary standards explains why we are inclined to consider her less criticizable than the weak-willed agent (see section 2).

One assumption of this analysis is that what makes particular actions compulsive is not the “strength” of their motivational antecedents.¹⁶ Strong desires and inclinations are felt by most normal persons from time to time, but they do not necessarily act on them. There is no reason to assume that these desires and inclinations are not occasionally as strong as the inclinations of a compulsive person, e.g., that a normal person’s desire and inclination to engage in extramarital sex cannot be of the same strength as a compulsive smoker’s inclination to light up another cigarette.

The difference between instances of compulsive and non-compulsive motivation is not so much a matter of their separate strengths as it is a matter of the way they operate in the person's psyche. Unlike non-compulsive motivation, compulsive inclinations are triggered directly by environmental cues via a process that is dissociated from the person's desires, deliberation, and, in many cases, even conscious beliefs. As a result of repetition and reinforcement they become entrenched. It is their frequency, cue-dependence, and dissociated nature, along with their disruptive impact on the normal psychological functioning of the person's deliberative and volitional processes, which set them apart and make them more difficult to control. While in some cases they might impair a person's capacity for rationality, such as her capacity to understand the facts involved in a decision, to weigh risks and benefits and evaluate consequences, in other cases they might impair her executive or volitional powers, such as her capacity to translate deliberative conclusions or decisions into action (most cases of compulsivity are likely to involve combinations of these effects).¹⁷ The proposal I now wish to entertain can be stated as follows:

- (Df.2) A particular action *A* is performed compulsively by (an agent) *S* if and only if *A* satisfies the following conditions: (i) *A* is part of a behavior pattern that *S* regularly performs in characteristic circumstances; (ii) *S* has inclinations to perform *As* which drive *S*'s performance of *As*; (iii) *S* would have continued to perform *As* despite (a) having desires to refrain from *As* and no desires to perform *As*, (b) judging that there are better reasons to refrain from *As*, and even (c) making sincere efforts to refrain from *As* which, other things held constant, would have been sufficient for a normal person to successfully refrain from *As* if she had equally strong desires and inclinations to perform *As*.

In contrast to Watson's account, the present proposal does not imply the loss of the compulsive persons' ability to resist at the time of action. Condition (iiic) introduces a restriction on the counterfactual characterizing the compulsive person's lack of control. It restricts this counterfactual to the range of possible worlds in which the effort she makes to refrain is sufficient relative to what is required of a *normal person* in these worlds to successfully refrain from actions of the same type if the normal person has equally strong desires and inclinations to perform them. But this effort will still be insufficient relative to what is required in these worlds of *her* to successfully refrain. If we assume a broadly hypothetical theory of ability, this is consistent with the compulsive person retaining her ability to refrain, since in a suitable context of relevantly similar possible worlds in which she invests greater efforts which are sufficient relative to what is required of *her* in these worlds, she will successfully refrain from the compulsive actions. Relying on the notion of "effort" in this way suggests that compulsion is a matter of degree. The greater the dissociation of the agent's inclinations to perform a type of action from her desires to perform actions of that type and the more entrenched these actions are, the greater the effort she will have to make in order to successfully refrain from actions of that type compared with what, other things held constant, would have been sufficient for a

normal person to successfully refrain from actions of the same type if she had equally strong desires and inclinations to perform them—and the more compulsive her actions are if she fails to refrain from them.¹⁸

Let me now sum up the present proposal and explain how compulsion differs from weakness of the will. What makes an action compulsive is that it is part of a habit that the compulsive person would find extremely difficult to discontinue even if she made a sincere effort to do so, where the reason for this difficulty is that the continuation of the habit counterfactually depends on a dissociation in her motivational system between desire and inclination, a dissociation which disrupts the normal psychological functioning of her deliberative and volitional processes. Now, some philosophers argue that the difference between compulsion and weakness of will is purely normative (Holton, 2009; Watson, 2004a). While the former is said to occur because the contrary motivation is too strong, the latter is said to occur because the will is too weak. However, this way of drawing the distinction is insufficient because compulsive actions can occur in the absence of contrary motivation. The present account provides a different way of drawing the distinction. Although it is true that the difference between compulsive and weak-willed actions is partly normative (as we have seen, both concepts contain normative elements), this is not the whole story: unlike the compulsive person, the weak-willed person would stop performing weak-willed actions if the desire to perform them evaporated and/or a sincere effort was made to refrain. That is because these actions are not part of a habit the continuation of which counterfactually depends on any motivational dissociation of the kind just described. Clearly, this is more than a normative difference. Indeed, given this metaphysical account of the notion of compulsive action, can we say addictive actions are compulsive? They are to the extent that addicts are inclined to perform them despite not desiring to do so, that is, to the extent that they carry on seeking and taking drugs even when they lose their desire to do so and make a sincere effort to refrain. Since there appears to be evidence that this is indeed the case (see section 3), the view that addictive actions are compulsive seems better supported than the view that they are weak-willed.

5. Objections with Replies

I want to end by considering some objections to this proposal. The first objection can be put like this: the notion of a “compulsive action” is analytically equivalent to that of a “motivationally *compelled* action,” i.e., an action the agent could not have refrained from performing due to irresistible motivation. Yet according to the present account, compulsive actions are claimed to be performed freely in the sense that it is within the agent’s powers to refrain from them. My response to this objection is to deny any analytic connection between the notions of “compulsive” and “motivationally compelled.” Rather, I maintain that compulsive actions need not be motivationally compelled. This follows from the reasonable demand on any metaphysical account of compulsive action that it be, at a very minimum, consistent

with the clinical-scientific notion of such action, together with evidence that compulsive persons are in fact perfectly capable of refraining from their compulsive actions.

Consider again obsessive-compulsive disorder. A successful treatment modality appears to be response prevention therapy (de Silva, 2006). The aim of response prevention therapy is to break the relationship between the various trigger situations which provoke the compulsive inclination, and the compulsive action, by repeatedly exposing the agent to different trigger situations but preventing her from performing the compulsive action. For example, a compulsive person whose washing rituals are a result of an obsession about being contaminated by dogs, may be instructed to pat a dog and then refrain from washing her hands or take a bath a given period of time afterwards. As the sessions are repeated, the period will be extended. Research shows that if this pattern is followed under each of the trigger situations, the cumulative effect is progressively less discomfort and inclination to engage in the compulsive behavior. On the assumption that compulsion entails inability to refrain, it is hard to explain how compulsive persons could engage in this kind of exposure therapy. How could they comply with instructions to delay their response to a trigger situation if they lacked this ability? It might be objected that the presence of ability, rather than a precondition of engaging in the therapy, is on the contrary a causal effect of this engagement. Thus the claim should be that compulsive persons have lost their abilities in ordinary circumstances which do not include therapy. That, of course, does not rule out the possession by these same agents of these abilities in counterfactual circumstances which do include therapy. The evidence is consistent with the possibility that the ability to refrain is initially absent but can be recovered as a result of complying with the therapist's instructions. Logically, that is, of course, true. But one would still have to explain *when* and *how* the ability is recovered. Is it recovered the moment the agent agrees to therapy? That seems odd. Why should this simple mental act cause the ability to refrain to pop up all of a sudden? Is it recovered when the agent delays her response to a trigger situation for the first time? That would clearly already require the presence of the ability. In general, it is difficult to see, given the assumption of inability, precisely how response prevention therapy can help compulsive persons regain control over their actions. If this is correct, inference to the best explanation would give preference to the view that these abilities were there all along, but compulsive persons face particularly tenacious difficulties in exercising them. What response prevention therapy does is to encourage these agents to put more effort into exercising abilities they already have.

The second objection I want to consider goes as follows: according to the present proposal, a compulsive inclination is a kind of simple intention similar to intentions guiding habitual behavior. As such, it does not work through the person's ordinary intentional system. If this is correct, compulsive behavior must be automatic, in which case the flexible and adaptable actions of, e.g., drug addicts cannot count as compulsive. In other words, the present proposal seems to imply a "by-passing" view of compulsivity according to which compulsive actions really are a kind of automatic behavior beyond the scope of conscious agency. In response to this objection,

consider the following: although compulsive inclinations are assumed to be a kind of simple intentions acquired as a result of conditioning independent of the person's ordinary intentional system, the details of the actions they explain need not be specifically represented in the contents of these intentions. That is, a typical intention may be expressible in the form of an imperative like "do drugs now!"¹⁹ An intention of this general kind can be executed in any number of ways, leaving the compulsive person with a lot of control over *how* she wants to realize it. In other words, what makes a *particular* action compulsive is not that it (necessarily) rules out flexibility, adaptability and decision-making. A particular action is compulsive if it is part of a habit the person would find extremely difficult to discontinue even if she made a sincere effort to do so. But if compulsive actions can be flexible and involve decision-making, one might object, how can they form parts of a *habit*? This objection depends on an all-or-none conception of automatic processes, a conception that has come under growing criticism in contemporary psychology where many see automaticity in terms of a gradual scale (Moors & De Houwer, 2006). According to the gradual view, a process may have both automatic and non-automatic features (Bargh, 1992). Habits seem a good example. They clearly have automatic features in the sense of their ability to be exercised efficiently and effortlessly, with minimal attentional resources. Yet habits also require the person to be sensitive to the circumstances in which she exercises them. Thus, normally, if her circumstances change, she will quickly respond and adapt her habitual actions (e.g., find a different route for her daily run if she has a running habit, switch to another brand of cigarettes if she has a smoking habit, etc.). It is not the case, therefore, that being part of a habit rules out either behavioral flexibility or, indeed, decision-making.

Finally, I want to consider the following objection: according to the present proposal, the fact that compulsive persons fail to refrain from acting upon their inclinations does not mean they could not have put more effort into refraining and, as a consequence, successfully have refrained. They have the ability to refrain. But assuming that they have this ability and that they do not suffer from weakness of will, why do they make insufficient effort? If they *want* to refrain, there simply is no way to explain this. It is more plausible, therefore, to assume that they have lost the ability. However, it is not the case that there are no ways of explaining why compulsive persons put insufficient effort into refraining without implying that they suffer from weakness of will. Let me end with some suggestions of possible explanations. Suppose, as some psychologists do, that self-control relies on an energy resource that comes in limited quantities and can be used up (Baumeister, Bratslavsky, Muraven, & Tice, 1998). Then attempting to refrain from acting upon continuing inclinations is likely to deplete these agents' self-control resources, making it increasingly harder for them to exercise their capacities successfully.²⁰ Since "increasingly harder" suggests a steady rise in the effort these agents need to invest in order to successfully refrain over time, there could well be a constant mismatch between their past experience of how much effort is required to refrain, and how much effort is required to refrain in the present. In combination with fatigue induced by depletion of self-control resources,

it might make them more likely to systematically underestimate the latter. Now, *misjudgment of effort* is surely a different fault from weakness of will (Watson, 2004a).

This might not be the only possible explanation. Another, related to a topic that has been much explored in social psychology, could be that low expectations of succeeding in their goals (e.g., abstaining from drugs) cause compulsive persons to make insufficient effort to refrain in spite of sincerely trying and independent of their actual abilities (Bandura, 1997). Thus, it seems possible that, following repeated failure to refrain, these agents may lose their sense of personal efficacy, and therefore unconsciously invest less effort in refraining from what they would have done, and what might have been sufficient relative to what is required, had they had higher expectations of a successful outcome. Assuming that the degree of compulsivity is determined by how much more effort these agents would have to shoulder in order to refrain compared with what would have been sufficient for normal persons, such systematic under-investments would surely make repeated failures to refrain quite likely. Once again, *unconscious underinvestment of effort* is a different fault from weakness of will.

There may be other explanations as well. But what they have in common is that they neither imply that compulsive persons have weak wills, nor do they rule out that they retain their abilities to control their compulsive actions and hence that these actions are freely performed.

6. Conclusion

Normative thinking about drug addiction has traditionally been divided between, on the one hand, a medical model which sees addiction as a brain disease characterized by compulsive drug-oriented behavior over which the addict has lost all control and, on the other, a moral model which considers drug addiction a moral weakness characterized by intentional drug-oriented actions under the full control of the addict. In this paper I have argued for a middle-ground position: although an essential feature of addictive actions is their compulsiveness, it is not the case that addicts have lost their ability to control these actions. That is because compulsive drug-oriented actions do not occur because of irresistible desires for drugs. They occur because they are parts of habits which are extremely hard to break. It is still within most addicts' powers to refrain from the addictive actions that form parts of these habits.

Acknowledgements

Many thanks to Helen Steward, Michael Morreau, Olav Gjelsvik, John Richard Sageng, Kari Poikolainen, Stephen Morse, Christel Fricke, Peter Railton and Richard Holton. Thanks also to two anonymous referees for this journal for very helpful comments. This paper was written with financial support from the The Research Council of Norway, Programme on Alcohol and Drug Research.

Notes

- [1] I leave open here whether ‘favorable’ should be a placeholder for some evaluative concept, such as DESIRABLE, PLEASURABLE, ENJOYABLE, or GOOD, or whether an evaluative neutral concept describing some positive affective state involving the object of the desire might be more appropriate. Different versions of the standard account can be found in the work of numerous authors. A common view is that it captures, roughly, what we mean by the notion of desire in ordinary language. For a recent alternative, however, see Schroeder (2004).
- [2] In one of the early experiments, about 70 percent of the voucher subjects were continuously abstinent for the first five weeks of the program, whereas fewer than 20 percent of the addicts in traditional psychological counseling maintained abstinence for the first five weeks (Heyman, 2009, p. 106).
- [3] Although I invoke the traditional understanding of weakness of will here as “akrasia,” nothing in what follows turns on such understanding. Similar arguments would apply to non-akratic weakness of will, i.e., unreasonable revisions of resolutions. For a defense of the latter view of weakness, see Holton (1999).
- [4] The claim that the weak-willed agent allows herself to give in to her recalcitrant desire by intentionally not making any sincere attempt to resist does not rule out any association of weakness of will with a sense of struggle. For example, it is consistent with the possibility that she makes half-hearted attempts to resist, even though she fails in the end to refrain from the weak-willed action. What it rules out is that she would have failed to refrain if her attempt at resisting had been *sincere*. This follows from the plausible assumption that the weak-willed agent is rationally criticizable because she could have refrained from her weak-willed action if she had made what by ordinary standards would count as a “sufficient effort,” something which was within her actual powers. For more on the notion of sufficient effort, see section 4.
- [5] For an objection, see Watson (2004a).
- [6] If weakness of will comes in degrees, why not say that this difference in attitudes arises because we are inclined to consider the non-addict experimenter as more weak-willed than the addict? I agree that weakness of will is likely to come in degrees. However, consider the following version of the example: suppose we knew that the addict and the non-addict experimenter made, roughly, the same efforts of will to refrain from acting upon their desires for heroin. Would we not still be inclined to find the addict less criticizable for her failure to refrain? My sense is that we would, which suggests that we are inclined to view the addict’s failure more as the result of her being subject to an inclination that is hard for her to control than as the result of having a too weak a will to control her inclination, and hence to refrain from acting upon it.
- [7] That addictive behavior tends to have these features are emphasized by several addiction researchers. See, e.g., Lowenstein (2000), and West (2006).
- [8] Perhaps the most striking example of addictive actions in the absence of positive value was reported by Lamb and colleagues (Lamb et al., 1991), who found that doses of drugs too low to produce any conscious experience at all may still activate implicit motivation to take drugs, as indicated by an increase in drug-seeking behavior. Thus, self-administration rates of morphine or cocaine injections which were too low to produce any subjective effects or positive appraisal, nevertheless significantly exceeded placebo injections in post-addicts.
- [9] The evidence here is plentiful and too comprehensive to be reviewed in this essay, spanning from studies on the neurobiology of sensitization to behavioral sensitization effects in humans from repeated drug exposure. For a review, see Robinson and Berridge (1998).
- [10] In fact, this way of distinguishing wanting from desiring is suggested by Berridge (2009) himself. Here he remarks that “wanting . . . may detach from the object of desire and be attributed too widely among stimuli, spewing indiscriminate ‘wanting’ in directions that are inappropriate or completely general” (Berridge, 2009, p. 388).

- [11] B&R (2003, p. 42) argue that while prefrontal and other cortical areas mediate cognitive forms of motivation, NAcc-related circuitry (in the nucleus accumbens), mediate Pavlovian-guided attributions of incentive salience. Addictive drugs are known to increase the concentration of dopamine in the NAcc in a number of ways.
- [12] Holton (2009) seems to be led by this to think that the motivational dissociation characteristic of addiction according to B&R should be interpreted in terms of a disconnection between addictive desire and evaluative judgment rather than between addictive inclination and desire. If the notion of “desire” is assumed to be as described in section 2, it is not clear whether Holton’s interpretation really captures the depth and radical nature of the motivational dissociation described by B&R. A further difficulty in assuming this interpretation is that it does not appear to clearly distinguish addictiveness from (akratic) weakness of will.
- [13] For a similar counterfactual account of control, see Fischer and Ravizza (1998).
- [14] This should not be confused with the technical way in which this notion is used by some philosophers of action.
- [15] Watson (2004a, pp. 56–57) makes a similar point.
- [16] There may be different views of what is meant by “strength” here. While some understand it in terms of the causal force of the person’s motivational states, others see it in terms of the intensity with which such states direct the person’s attention to their objects in terms of some evaluative category (Wallace, 2006; see also section 2). For present purposes, it does not matter which of these views is assumed. The essential difference between instances of compulsive and non-compulsive motivation is neither their causal force nor the intensity with which they direct the person’s attention to the pleasurable features of their objects (which does not rule out, of course, that compulsive inclinations often have a lot of causal force or often are associated with a focus on the pleasurable features of their objects). See below.
- [17] Many scientists think that addiction essentially involves failures in decision-making arising from the interactions between a slow-learning habit system and a fast-learning planning system. For a detailed review of some of these failures, see Redish, Jensen, and Johnson (2008).
- [18] Compulsivity, it is important to note, is a matter of degree. The account is consistent, therefore, with the possibility of differences between forms of compulsive actions in terms of how difficult they are to resist, e.g., between the kinds of actions seen in, say, OCD and in addiction.
- [19] For the idea that inclinations have this form, see Schapiro (2009).
- [20] Several philosophers have suggested that “ego-depletion,” which this phenomenon is called, may play a part in compulsive or addictive behavior. See, e.g., Levy (2006) and Zaragoza (2006).

References

- Abramowitz, J.S., & Deacon, B.J. (2005). Obsessive-compulsive disorder: Essential phenomenology and overlap with other anxiety disorders. In J.S. Abramowitz & A.C. Houts (Eds.), *Concepts and controversies in obsessive-compulsive disorder* (pp. 119–133). New York: Springer.
- Austin, J.L. (1979). A plea for excuses. In his *Philosophical papers* (pp. 175–204). Oxford: Clarendon Press.
- Bandura, A. (1997). *Self-efficacy, the exercise of control*. New York: W. H. Freeman and Company.
- Bargh, J.A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology*, 105, 181–199.
- Baumeister, R.F., Bratslavsky, E., Muraven, M., & Tice, D.M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.

- Benn, P. (2007). Disease, addiction and the freedom to resist. *Philosophical Papers*, 36, 465–481.
- Berridge, K.C. (2009). Wanting and liking: Observations from the neuroscience and psychology laboratory. *Inquiry*, 52, 378–398.
- Davidson, D. (1980). How is weakness of the will possible? In his *Essays on actions & events* (pp. 21–42). Oxford: Clarendon Press.
- de Silva, P. (2003). The phenomenology of obsessive-compulsive disorder. In R.G. Menzies & P. de Silva (Eds.), *Obsessive-Compulsive disorder, theory, research and treatment* (pp. 21–38). Chichester, UK: John Wiley & Sons.
- de Silva, P. (2006). Obsessive-Compulsive disorder. In A. Carr & M. McNulty (Eds.), *The handbook of adult clinical psychology* (pp. 487–522). New York: Routledge.
- Evans, J., & Coventry, K. (2006). A dual-process approach to behavioral addiction: The case of gambling. In R.W. Wiers & A.W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 29–44). Thousand Oaks, CA: Sage.
- Fischer, J.M., & Ravizza, M. (1998). *Responsibility and control: An essay on moral responsibility*. Cambridge: Cambridge University Press.
- Frankfurt, H. (1988). Freedom of the will and the concept of a person. In *The importance of what we care about* (pp. 11–26). New York: Cambridge University Press.
- Heather, N., Tebbutt, J.S., Mattick, R.P., & Zamir, R. (1993). Development of a scale for measuring impaired control over alcohol consumption: A preliminary report. *Journal of Studies on Alcohol*, 54, 700–709.
- Heyman, G.M. (2009). *Addiction, a disorder of choice*. Cambridge, MA: Harvard University Press.
- Holton, R. (1999). Intention and weakness of will. *The Journal of Philosophy*, 96, 241–262.
- Holton, R. (2009). *Willing, wanting, waiting*. New York: Oxford University Press.
- Lamb, R.J., Preston, K.L., Schindler, C.W., Meisch, R.A., Davis, F., Katz, J.L., . . . Goldberg, S.R. (1991). The reinforcing and subjective effects of morphine in post-addicts: A dose-response study. *Journal of Pharmacology and Experimental Therapeutics*, 259, 1165–1173.
- Levy, N. (2006). Addiction, autonomy and ego-depletion: A response to Bennett Foddy and Julian Savulescu. *Bioethics*, 20, 16–20.
- Lowenstein, G. (2000). Willpower: A decision theorist's perspective. *Law and Philosophy*, 19, 51–76.
- Lyvers, M. (2000). Loss of control in alcoholism and drug addiction. *Experimental and Clinical Psychopharmacology*, 8, 225–249.
- Mele, A. (1990). Irresistible desires. *Nous*, 24, 455–471.
- Mele, A. (2009). *Effective intentions: The power of conscious will*. New York: Oxford University Press.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132, 297–326.
- Morse, S.J. (2000). Hooked on hype: Addiction and responsibility. *Law and Philosophy*, 19, 3–49.
- Pollard, B. (2006). Explaining actions with habits. *American Philosophical Quarterly*, 1, 57–69.
- Quinn, W. (1995). Putting rationality in its place. In R. Hursthouse, G. Lawrence, & W. Quinn (Eds.), *Virtues and reasons: Philippa Foot and moral theory* (pp. 181–208). New York: Oxford University Press.
- Redish, A.D., Jensen, S., & Johnson, A. (2008). A unified framework for addiction: Vulnerabilities in the decision process. *Behavioral and Brain Sciences*, 31, 415–487.
- Robinson, T.E., & Berridge, K.C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*, 18, 252–255.
- Robinson, T.E., & Berridge, K.C. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, 28, 309–369.
- Robinson, T.E., & Berridge, K.C. (2003). Addiction. *Annual Review of Psychology*, 54, 25–53.
- Schapiro, T. (2009). The nature of inclination. *Ethics*, 119, 229–256.
- Schroeder, T. (2004). *Three faces of desire*. New York: Oxford University Press.
- Shoemaker, S. (1996). On knowing one's own mind. In his *The first person perspective and other essays* (pp. 25–49). Cambridge: Cambridge University Press.

- Smith, M. (2004). Rational capacities, or: How to distinguish recklessness, weakness, and compulsion. In S. Stroud & C. Tappolet (Eds.), *Weakness of will and practical irrationality* (pp. 17–38). New York: Oxford University Press.
- Steward, H. (2009). Animal agency. *Inquiry*, 52, 217–231.
- Wallace, R.J. (2006). Addiction as defect of the will: Some philosophical reflections. *Normativity & the will* (pp. 165–189). New York: Oxford University Press.
- Watson, G. (2004a). Skepticism about weakness of the will. *Agency and answerability* (pp. 33–58). New York: Oxford University Press.
- Watson, G. (2004b). Disordered appetites: Addiction, compulsion, and dependence. In his *Agency and answerability* (pp. 59–87). New York: Oxford University Press.
- West, R. (2006). *Theory of addiction*. Oxford: Blackwell.
- Zaragoza, K. (2006). What happens when someone acts compulsively. *Philosophical Studies*, 131, 251–268.