INTENTION AND WEAKNESS OF WILL*

There is something curious about the philosophical literature on weakness of will. It is not about what one might expect. Even David Wiggins,[1] in a discussion that has much in common with that to be given here, starts out by claiming:

> Almost anyone not under the influence of theory will say that, when a person is weak-willed, he intentionally chooses that which he knows or believes to be the worse course of action when he could choose the better course (*ibid.*, p. 239).

I do not agree that this is the untutored view. Whenever I have asked nonphilosophers what they take weakness of will to consist in, they have made no mention of judgments about the better or worse course of action. Rather, they have said things like this: weak-willed people are irresolute; they do not persist in their intentions; they are too easily deflected from the path that they have chosen. My aim here is to pursue this line of thought. I shall develop the idea that the central cases of weakness of will are best characterized not as cases in which people act against their better judgment, but as cases in which they fail to act on their intentions. Of course, to say this is not to say enough. Not every case of a failure to act on one's intentions is a case of weakness of will. Sometimes, we realize that our intentions were ill-judged, or that circumstances have changed to make them inappropriate. When, in such cases, we fail to act on our intentions, we need not display weakness of will. So a large part of the discussion below will be concerned with saying what kind of failure to act on one's intention does constitute weakness of will. The basic idea is this. Just as it takes skill to form and employ beliefs, so it takes skill to form and employ intentions. In particular, if intentions are to fulfil their function, they need to be relatively resistant to reconsideration. Weakness of will arises, I shall suggest, when agents are too ready to reconsider their intentions.

* This paper was written while I was a fellow at the Research School of Social Sciences, Australian National University. Thanks to audiences there, and at Monash University, Sydney University, and Sheffield University for discussion; and especially to Sam Guttenplan, Lloyd Humberstone, Rae Langton, David Lewis, Philip Pettit, and Michael Smith.

[1] "Weakness of Will, Commensurability, and the Objects of Deliberation and Desire," in his *Needs, Values, Truth* (Cambridge: Basil Blackwell, 1987), pp. 239-67.

0022-362X/99/9605/241–62

In taking this approach, I depart from almost all of the literature on the subject.[2] Since I find the approach obvious, this calls for some explanation. I suspect that there are two factors at work. First, there has been, until quite recently, a widespread suspicion of the notion of intention. Some have said that intentions are a philosopher's invention, a misleading reification of the legitimate talk of intentional action. Others, while prepared to concede a place to intentions, have held them to be reducible to beliefs and desires, or else to beliefs and desires and their relations to action. Such considerations have led many philosophers to think that an account of weakness of will must ultimately be couched in the language of beliefs and desires. The traditional understanding of weakness of will as failure to do what one judges to be best is couched in this language; so, too, are the main rival accounts.[3] In the last few years, however, a number of philosophers have argued that talk of intentions is both legitimate and irreducible: I have in mind especially the work of Michael Bratman and Gilbert Harman.[4] If they are right that such talk is legitimate, then there is no *requirement* to analyze weakness of will in terms of belief and desire rather than in terms of intention. That is all that

[2] Exceptions: Wiggins (*ibid.*) goes on to sketch an account of weakness of will that does tie it to failure to persist in one's intentions. Alfred Mele—*Irrationality* (New York: Oxford, 1987)—discusses cases of weakness of will in which one acts against one's intention, but he still treats this as a special (and specially problematic) case of action against one's best judgment; that is, these are cases in which one judges a particular course as right, and then forms the intention to pursue it, and then fails to act on it. In contrast, I suggest that action against one's judgments about what is right need not enter into examples of weakness of will at all. Amelie Rorty insists in various articles that weakness of will should not be identified with *akrasia*; she also charts a number of places at which the "akratic break" can take place, including at the point between forming an intention and acting on it—"Where Does the Akratic Break Take Place?" *Australasian Journal of Philosophy*, LVIII (1980): 333-46. I discuss some of the issues raised by her paper below. Christopher Peacocke argues that an account of intention is essential to an account of *akrasia*—"Intention and *Akrasia*," in Bruce Vermazen and Merrill B. Hintikka, eds., *Essays on Davidson: Action and Events* (New York: Oxford, 1985), pp. 51-73; but his project is very different from that offered here. Roughly, he is concerned with the idea that *akratic* action results from *forming an intention* (instead of simply performing an action) that is not in line with one's judgment as to what is best.

[3] Namely, accounts that analyze weakness of will in terms of a failure to bring one's lower-order desires into line with one's higher order desires, as presented in John Bigelow, Susan Dodds, and Robert Pargetter, "Temptation and the Will," *American Philosophical Quarterly*, XXVII (1990): 39-49, following Richard C. Jeffrey, "Preference among Preferences," this JOURNAL, LXXI, 13 (July 18, 1974): 377-91; and accounts that analyze weakness of will in terms of desire revision that does not result from conditionalization, as presented in Frank Jackson, "Weakness of Will," *Mind*, XCIII (1984): 1-18.

[4] Bratman, *Intention, Plans, and Practical Reason* (Cambridge: Harvard, 1987); Harman, *Change in View* (Cambridge: MIT, 1986).

is needed for the account offered here. If, in addition, they are right that talk of intentions is irreducible, then it will not be possible to rephrase the account offered here in terms of beliefs and desires. This, however, is an issue that I shall not broach; although I am sympathetic to the irreducibility claim, I shall not presuppose it in what follows.

The second factor which has influenced the lines of the discussion is the impetus given by Plato and Aristotle. Both clearly were concerned with the question of whether it is possible to act against one's best judgment; and much of the contemporary literature has been concerned with assessing and responding to their arguments. Let us grant the term *'akrasia'* to refer to action voluntarily undertaken against one's best judgment; it is, after all, scarcely a term of ordinary English. Then my contention will be that weakness of will is not *akrasia.*

The outline of my discussion is as follows. In section I, I give an account of intentions, relying largely on Bratman's work, but going beyond him in some areas. I show in section II how this leads naturally to an account of weakness of will. In section III, I chart the considerable advantages that this account has over the traditional account that identifies weakness of will with *akrasia.* I reply to some possible objections in section IV.

I. INTENTIONS

Bratman's account of intentions has three parts. He identifies the characteristics that a state needs to have if we are to count it as an intention; he explains the useful role that states with these characteristics would have in our lives; and he gives reasons for thinking that beliefs and desires on their own could not fulfill this role. Taken together, these provide the basis for the case that we should acknowledge the existence of intentions. For if we seem to have them, and if having them would play a useful role, and if indeed they do seem to play this useful role, then the burden of proof rests in their favor. We shall need compelling arguments before giving them up.

I shall not be concerned here with the third element of Bratman's argument: the contention that beliefs and desires cannot play the role allotted to intentions. Readers in search of that can read Bratman's book. But the other two elements of his argument—the account of the distinguishing marks of intentions, and of the useful role that intentions can play—will be central to my account of weakness of will. Let me start by identifying the distinguishing marks; in particular, let me start with the distinguishing marks of *future-directed*

intentions (that is, intentions to do something at some future time). I shall return to consider present-directed intentions later.

A future-directed intention to perform a certain action typically has the following characteristics. The agent forms the intention at one time by making a decision to perform the action. Then, unless it is revised, the intention will directly lead the agent to perform the action; it is, as Bratman says, *controlling* (*op. cit.*, p. 16). Moreover, it is relatively immune to reconsideration and hence to revision. Once formed, intentions have a tendency to persist. They have what Bratman calls *stability* (*op. cit.* pp. 16, 65).

What need do we have of states that are stable and controlling in this way? Part of Bratman's answer stems from the observation that we are epistemically limited creatures. Information is scarce and costly to obtain. Reasoning on the information that we have takes time and effort. It is rational then to limit the amount of time we spend looking for information and reasoning on it. We should do a certain amount of searching and reasoning, and then stop and make a decision on the basis of that. Now, if we were to act immediately on making the decision, then this would not in itself give rise to a need for intentions. But suppose our decision is a decision to act tomorrow. Then we need some way of storing our decision so that we act on it tomorrow without reconsidering it (for to reconsider it would violate the requirement that we not reason about it further). In short, we need a state that is stable and controlling: an intention to perform the action that is relatively immune to revision, and will in itself be enough to move us when the time comes.

This, in turn, raises another question: Why should we want to make a decision now about how we shall act tomorrow? It could be that now is a more propitious moment to reason. Tomorrow, there will be some factor that will stop us from reasoning well: we shall be short of time, or distracted, or some such. But a more common reason for doing our reasoning ahead of time is that all sorts of other actions will be dependent upon what we decide to do; and we shall need to perform some of these actions in the meanwhile. Thus, suppose I want to paint my front door tomorrow. Then there might be many reasons for deciding today what color to paint it. It could be that tomorrow I shall start at the crack of dawn when there will be too little light to see the color charts properly; that will give me a reason of the first kind for doing my deliberation today. More probably, the paint shop will be closed tomorrow, so I should buy the paint today; and since I want only to buy the color that I shall use, I shall need to have decided today what color that will be. I need to form

my intention ahead of time. Thus, intentions are necessary in the *intrapersonal* case for agents to coordinate their activities. They are also important in the *interpersonal* case. Suppose some friends are coming to visit me tomorrow, and the easiest way for them to recognize my house is from the color of the front door. I shall be speaking with them today, but not again before they visit. Then, once again, I have a reason for forming today an intention about which color I shall paint my door tomorrow, so that I can tell them what to look for.

It will be important to my argument to realize that there can be good reason for forming intentions now, even in cases where I cannot make my decision just on the basis of rational deliberation. One such case involves Buridan-style examples: if I am really indifferent between red and green for my door, it is nonetheless important that I decide on one of them, and form the corresponding intention. Otherwise, I shall not know what paint to buy, or what to tell my visitors. A more interesting case concerns options that the agent finds incommensurable. Suppose (taking an example at random) that I do not know how to compare the demands of leaving home to fight fascism with the demands of staying to look after my mother. Here, the problem is not that each outcome fares equally well on the same scale; I do not know how to place them on the same scale. Nevertheless, I need to decide on one course, and to form the corresponding intention; both intrapersonal and interpersonal coordination require this. I need to know whether to start winding up my affairs and packing my bag; and I need to know what to tell my mother so that she too can make her plans accordingly.[5]

I have sketched the main lines of Bratman's argument for the utility of intentions. Once we have got this far, it is easy to see that they might be useful in other roles, too, roles that Bratman does not discuss. Suppose that there is an act that I now believe I should perform at some time in the future; and suppose I know that when the time comes, I shall not want to perform it. Then it would be useful to form an intention now, an intention which will lead me directly to act when the time comes, and which will be somewhat immune to reconsideration in the light of the desires I shall have then. Philosophers have been keen on

[5] Bratman points out that cases like these show the utility of intentions even in cases where information is not limited; for we have reason to think that Buridan cases will arise even for agents who know everything. He also argues, I think convincingly, that they show that we cannot reduce talk of intentions to talk of beliefs and desires. By hypothesis, I do not desire that I take one course rather than the other; and my intention to take one course cannot be seen as a belief that I will take it. See Bratman, "Davidson's Theory of Intention," in *Essays on Davidson: Actions and Events*, pp. 13-26, here pp. 22-23; *Intention, Plans, and Practical Reason*, pp. 11, 22ff.

the idea that sometimes an intention will not be enough; that we will need to bind ourselves in some further way to fulfil our intentions, perhaps by placing temptation out of our reach, or by telling others of our plans so that fear of their disapproval should we fall short will provide an extra incentive.[6] But we should not let our interest in these exotic methods blind us to the fact that very often, intention is enough. People get up on cold dark mornings, leave enjoyable lunches to return to work, even give up smoking, using no mechanism other than an intention to overcome their contrary desires.[7]

The use of intention to overcome present desire becomes especially important in cases that involve repeated actions. I reason that since smoking forty cigarettes a day for the rest of my life will make a considerable difference to my chance of getting lung cancer, I should give it up. But should I deny myself the cigarette I was about to have? Smoking one cigarette will make virtually no difference to my chances of getting lung cancer. So why should I deny myself? Unfortunately, the same argument will work just as well forty times a day for the rest of my life. What we need here is not just an intention to perform—or, in this case, to refrain from performing—a specific action. We need a general intention concerning a certain type of action; what Bratman calls a *policy*,[8] and I shall call a *policy intention.*

To conclude this section, let us return to an issue shelved at its beginning: Are there such things as *present-directed* intentions? All the considerations raised so far have concerned future-directed intentions, that is, intentions to do something in the future. Do they give us any reason for accepting intentions to perform actions straight away? I rather doubt that they do; however, there is a slight theoretical elegance in accepting present-directed intentions. I spoke of future-directed intentions as resulting from a decision to act. I do not mean by this that there are two distinct states: the decision and the intention. Rather, to form the intention just is to make the decision.

[6] Jon Elster, *Ulysses and the Sirens* (New York: Cambridge, 1979). I suspect that very often the point of telling others of our intentions is not to incur their scorn if we revise them. Rather, it is simply to remind us of what our intentions were, or of how seriously we made them. A midwife told me of the entry on one woman's birth plan: "If I ask for pain relief in labor, show me this birth plan." The woman did ask for pain relief, the midwife showed her the birth plan, and the woman decided against taking any.

[7] Bratman briefly mentions examples in which intentions are used to overcome future resistance, but argues that we get a better understanding of the central cases of intention by ignoring them—*Intention, Plans, and Practical Reason*, p. 12. I suspect that such examples do provide central cases of intention; what we do well to ignore are examples in which intention is not enough.

[8] *Intention, Plans, and Practical Reason*, pp. 87-91.

(This is not to say that you cannot *have* an intention without having made a decision. You might have been born with the intention; or it might have been implanted by some manipulator using fiendish methods. But if so, you did not *form* the intention.) Then we might want to say that a decision to act now also results in an intention, but a present-directed one. A number of theorists have taken this path.[9] It seems to me to be slightly at odds with ordinary usage, but innocuously so; and it will make my account simpler if I adopt it here. So when I speak of intentions, these can be either present- or future-directed.

II. WEAKNESS OF WILL

I have argued that we have a need for intentions, that is, that we have a need for action-guiding states that are not readily revised. My suggestion now is that this gives the basis for an account of weakness of will. A person exhibits weakness of will when she revises her intention, in circumstances in which she should not have revised it. This 'should' is not meant in a moral sense. Rather, it is the 'should' that is generated by the norms of the skill of managing one's intentions. A person is weak willed if she revises her intentions too readily.

What is it to be too ready to revise an intention? Clearly, not every case of a revision is a case of an over-ready revision. If you and I intend to go for a picnic, and it starts to rain, we might be quite right to revise our intention, and in doing so we need not show weakness of will. (I only say we 'need not' show weakness of will in such a case, for I have left it underdescribed; if we had vowed to go whatever the weather, then we probably would have been weak willed in giving up when it rained.) It might be thought that we can accommodate cases like these by invoking conditional intentions. Even if we did not say so explicitly, it might be thought that our intention was the conditional one of going for a picnic provided it did not rain. But not every reasonable revision of an intention can be understood as the triggering of an escape clause in a conditional intention. We would be equally right to abandon our intention in the face of a freak plague of frogs falling from the sky; but neither of us considered that, even implicitly. We need, then, some general account of when it is right to revise our intentions.

Friedrich Nietzsche[10] seems to have held that it is *always* wrong to revise an intention:

[9] See, for instance, the notion of proximal intentions in Mele, *Springs of Action* (New York: Oxford, 1992), part II.

[10] *Beyond Good and Evil* (Harmondsworth: Penguin, 1973); cited also in Gary Watson, "Skepticism about Weakness of Will," *Philosophical Review*, LXXXVI (1977): 316-39, here p. 328.

> To close your ears to even the best arguments once the decision has been made: sign of a strong character. Thus an occasional will to stupidity (*ibid.*, §107).

Perhaps this is good advice for football referees. But as a general attitude it is surely a recipe for disaster. Moreover, it is not even true that such intransigence is a sign of a strong character. One is reminded of Somerset Maugham's[11] jibe against the Vicar of Blackstable: "Like all weak men he laid an exaggerated stress on not changing one's mind" (*ibid.*, chapter 39). Nevertheless, if our account of intentions is correct, there is something right about the idea that there should be a real reluctance to revise an intention. To see this, let us return once more to Bratman, and to his account of the circumstances in which we rationally should reconsider our intentions. His idea is that it is rational to reconsider an intention just in case doing so manifests tendencies that it is reasonable for the agent to have; similarly, it is rational to fail to reconsider an intention just in case this manifests tendencies that it is reasonable for the agent to have.[12] This means, of course, that sometimes it will be rational not to reconsider an intention even when reconsideration would, in that instance, have been in the agent's interests. In general, once we have decided in which restaurant to eat, it is a good idea to let the matter rest, without endlessly discussing the pros and cons; and this is true, even though it occasionally means that we shall go to a restaurant that we could have realized was not the best choice.

Presumably, the converse of Bratman's criterion also holds: it is not rational to reconsider something if the reconsideration manifests tendencies that it is not reasonable for the agent to have. Adding this to the account proposed here, we arrive at the following: actors show weakness of will when they revise an intention as a result of a reconsideration that they should not have performed; that is, when their reconsideration exhibits tendencies that it is not reasonable for the agent to have.

[11] *Of Human Bondage*(New York: Doran, 1915). Note that Maugham does not ascribe weakness of will to the vicar. There is more than one kind of weakness; what is at issue here is more like weakness of character. For a further distinction, consider J. L. Austin's complaint about the "grotesque confusion of moral weakness with weakness of will"—"A Plea for Excuses," in his *Philosophical Papers* (New York: Oxford, 1961), pp. 123-52, here p. 146, note 1.

[12] Bratman, *Intention, Plans, and Practical Reason*, p. 68. He actually talks in terms of habits rather than tendencies. But for something to be a habit, it must be habitually exhibited. In contrast, it is possible for a tendency to be only rarely exhibited, if at all. I want the latter notion.

As it stands, this will seem horribly vague. Which tendencies of reconsideration is it reasonable for an agent to have? I would suggest rules of thumb like the following:

(1) It is reasonable to have a tendency to reconsider intentions, if one believes that circumstances have changed in such a way that they defeat the purpose of having the intention.
(2) It is reasonable to have a tendency to reconsider intentions, if one believes that they can no longer be carried out.[13]
(3) It is reasonable to have a tendency to reconsider intentions, if one believes that they will lead one to great suffering when that suffering was not envisaged at the time of forming the intention.
(4) It is reasonable to have a tendency not to reconsider intentions in circumstances that prevent clear thought, if those intentions were made in circumstances that allow clear thought.
(5) It is reasonable to have a tendency not to reconsider intentions that were expressly made in order to get over one's later reluctance to act.

This is still pretty vague: the principles themselves are vague, they are doubtless incomplete, and I have said nothing about which of them should take precedence over which. But I shall not try to eliminate this vagueness. Vagueness is a problem when we try to determine the extension of a concept: when we try to provide a means of telling for each particular act whether it does or does not display weakness of will. My concern here is rather with giving an analysis of the concept itself. If the concept is vague, then the analysis had better be vague, too, and along just the same dimensions. I hope that we shall see that the account proposed here provides just the right sort of vagueness.

The account is not quite plausible as it stands. I need to make one clarification and one amendment. First, the clarification. Suppose I intend to perform a rash and dangerous act: leaping from a high cliff on an untested homemade hang-glider. Would I show weakness of will in revising my intention? That rather depends on why I revise it. If I reassess the chances that the hang-glider will fail, and decide on a program of more cautious preliminary testing, that is surely a reasonable revision. If, on the other hand, I simply suffer a failure of

[13] Perhaps this rule is otiose. Perhaps it is a necessary condition on the *formation* of an intention that one does not believe that one will be unsuccessful in carrying it out. Or, more controversially, perhaps one must actually believe that one will be successful. (For discussion and references, see Bratman, *Intention, Plans, and Practical Reason*, pp. 35–41; he denies that these are necessary conditions.) Equally then, perhaps it is impossible to *maintain* an intention in the face of a new belief that one cannot carry it out, or even in face of the loss of the belief that one can.

nerve, backing off from the edge of the cliff at the sight of the drop before me, a failure I would have suffered no matter how good the hang-glider, then that is plausibly a case of weakness of will. So the question is not simply whether or not it is reasonable to reconsider my intention; it is whether or not the particular reconsideration I perform is one that exhibits tendencies that it is reasonable for me to have.

Second, the amendment.[14] Some people go in for a lot of low-level intention revision. They decide to go to one restaurant, change their minds in favor of another, switch back to the first, again to the second, and so on. Now, it is not reasonable to have such tendencies; much time and effort is wasted, and interpersonal coordination becomes a nightmare. But while we might describe such people as fickle or capricious, we would not normally describe them as weak willed. So the account offered so far is too broad. We need to know what is distinctive about weakness of will, what it is that distinguishes it from caprice.

I mentioned before that one reason for forming an intention can be in an attempt to overcome contrary desires that one believes one will have when the time comes to act. An intention can be, as we might say, *contrary inclination defeating*. Now, this might be only part of the reason for forming a particular intention, and not a very important part at that. Indeed, sometimes this aspect might be merely implicit in an intention, and might only emerge if we asked the person what they would do if they were to change their preferences. The notion is not vacuous, however, since some intentions clearly are not contrary inclination defeating. Thus, someone might form an intention that has the content: I will go to The Red Lion tonight, providing I still feel like going (and such an intention can still have an important role in coordination, despite its conditional nature).

So we have a distinction between intentions which are designed to be contrary inclination defeating, and those which are not. This provides us with the grounds that we need for distinguishing weakness of will from caprice. If someone overreadily revises an intention that is, at least partially, contrary inclination defeating, that is weakness of will; if they overreadily revise an intention that is not, that is caprice. Consider again the vacillating diner. Suppose he has become concerned about his tendency to keep changing his mind, and so resolves to go to a particular restaurant even if another seems more

[14] Jeanette Kennett showed me that I needed to make an amendment here. Lloyd Humberstone showed me how to make it.

attractive later on. In other words, suppose he forms a contrary-inclination-defeating intention to go to that restaurant. Then if he revises his intention once again, he would not merely be capricious; he would display weakness of will.

Thus, we can think in terms of a genus with two species. The genus is irresoluteness: unreasonable revision of intention. The two species, weakness of will and caprice, are distinguished by whether the intentions revised are contrary-inclination-defeating intentions or not. I am not sure how theoretically important this distinction is; moreover, we have now introduced further vagueness at the boundary between the two species. But the distinction, and the vagueness at the boundary, does, I think, correspond to our ordinary usage. Since this is a paper on weakness of will, I shall limit discussion to intentions that are contrary inclination defeating; but since that is rather a mouthful, I shall just call them *intentions.*

III. ADVANTAGES OF THE PRESENT PROPOSAL

I shall present six reasons for believing the account of weakness of will which is offered here; or, to be more precise, six reasons for preferring it to the traditional account that understands weakness of will as a failure to do what one judges to be best.[15]

(*i*) *Accommodating weakness of will in cases of indifference or incommensurability.* I mentioned before that we form intentions in cases where we are indifferent between the alternatives, or treat them as incommensurable. When we do so, and revise our intentions, we can show weakness of will. That is just what the account proposed here would predict. But the traditional account can make no sense of such cases; for, by hypothesis, the agents involved do not judge one option to be better than the other, and so cannot be acting against their better judgment.

To see the possibility of such cases, consider again the example of incommensurability mentioned above. I am caught between the conflicting demands of fighting fascism and of staying to look after my mother. Unable to compare them in any meaningful way yet knowing I must choose, I decide to go and fight. But, then, when the moment to leave comes, I revise my intention and stay with my mother. That can be an example of weakness of will. (Note that to generate such a case we need not say that the options really are incommensurable; only that the agent judges them to be so.)

(ii) Explaining the relation of weakness of will to strength of will. One would expect the property opposed to weakness of will to be

[15] Some of these considerations also, I believe, provide reasons for preferring it to the accounts offered by Bigelow, Dodds, and Pargetter, and by Jackson; but I shall not argue for this contention here.

strength of will. And strength of will is something that shades off into stubbornness. The account I propose explains this in a straightforward way. A person shows strength of will when they stick by their intentions in circumstances in which they are right to do so; that is, when they do not reconsider them, and not doing so exhibits tendencies that it is reasonable for them to have. Strength of will turns to stubbornness when they stick by their intentions even when it is reasonable to reconsider and revise them. And, of course, the boundary between these two is vague; reasonable people will disagree about which tendencies of intention revision it is reasonable to have, just as they will disagree on when strength of will moves over into stubbornness. (One way of understanding the quotation from Nietzsche is to see him placing the boundary at one extreme.) Here, then, is one place where the vagueness in the *analysans* mirrors the vagueness in the *analysandum*.

In contrast, the traditional account has real problems in explaining strength of will. Indeed, defenders of the traditional account have typically contrasted weakness of will with self-control, where this is understood as the ability to do what one believes to be best. But self-control is not continuous with stubbornness, in the way that strength of will is. It is perhaps possible to have too much self-control; but the worry there is that one lacks spontaneity; and lacking spontaneity is very different from being stubborn.

(iii) Accounting for cases of oscillating weakness of will. Consider this example from Thomas Schelling[16]:

> As a boy I saw a movie about Admiral Byrd's Antarctic expedition and was impressed that as a boy he had gone outdoors in shirtsleeves to toughen himself up against the cold. I resolved to go to bed at night with one blanket too few. That decision to go to bed minus one blanket was made by a warm boy. Another boy awoke cold in the night, too cold to retrieve the blanket and resolving to restore it tomorrow. But the next bedtime it was the warm boy again, dreaming of Antarctica, who got to make the decision. And he always did it again (*ibid.*, p. 59).

Let us suppose that the cold Schelling had at least tried to retrieve the blanket. How would we describe the situation? Would we say that Schelling displayed weakness of will when he reached for the blanket despite his earlier resolution to leave it off? Or would we say that he displayed weakness of will in leaving the blanket off again the following night, despite his resolution to leave it on? I think that we might

[16] "The Intimate Contest for Self-Command," in his *Choice and Consequence* (Cambridge: Harvard, 1984), pp. 57-82.

well want to say both (we would need to fill out the details in the right way; he must form each intention seriously, having given it proper consideration, and so on). The account I am offering enables us to do so. We say that in each case, in overturning his previous intention, Schelling displays tendencies that it is not reasonable for him to have. But on the traditional account of weakness of will, it is not easy to see how we can get this result.

On the traditional account, the normal case is one in which the agent maintains a view about which action is the right one, and then acts contrary to this belief. But that cannot be the case here, since if the young Schelling maintained a view, then it would have been either that it is best to leave the blanket on, or that it is best to take it off. But if he maintained either of these views, then it is not possible for *both* actions to display weakness of will, since it is not possible for both actions to be contrary to the view maintained. Alternatively, proponents of the traditional account might think that the young Schelling changed his mind about which course was the best. Then, in order to characterize him as doubly weak willed, they would have to say that in the middle of the night, shivering beneath one blanket too few, he believed that what was best was to keep the blanket off, and his weakness stemmed from reaching for it; and they would have to say that when going to bed, warm and dreaming of the Antarctic, he believed that it was best to keep the blanket on, and his weakness stemmed from his leaving it off. But that is completely implausible; if there were any reason for thinking that, in the middle of the night, he believed it best to keep the blanket off, that would be because he believed this when he went to bed, and had not changed his mind; and similarly, if there were any reason for thinking that, at bedtime, he believed it best to keep the blanket on, that would be because he believed this in the middle of the night, and had not changed his mind.

(iv) Allocating the stigma of weakness of will. There is a considerable stigma attached to being weak willed. I think that the account offered here correctly predicts the times that it is deserved, and the traditional account does not. I have a friend who believes that all the arguments point to the same conclusion: he should not eat meat. But he is not moved. "I am," he says disarmingly, "inconsistent." Now there is something very odd about that, but, although he is *akratic*, he does not attract the stigma that attaches to weakness of will. Indeed, I would never call this friend weak willed. Suppose, however, that he announces that he has decided to give up meat; his intention is firm. If I then find him roasting himself a suckling pig in a pit, I

shall say that he is weak willed, and quite rightly. It is the failure to persist in the intention that makes all the difference.

We can make the point in another way. Suppose my friend tells me that he will give up eating meat on the first day of January; it is his New Year's resolution. I find him ruefully eating sausages on New Year's Eve; I cannot scorn him for his weakness of will. Yet were I to find him eating the same meal the following day, I could. His views about what is best have not changed over the two days; the difference stems from the intention.

These considerations are relevant to one of the binding strategies that was mentioned before. If I fear that I shall not persist in my intention, then one thing I can do is to make it public; for the scorn that I shall suffer if I revise the intention will provide me with an incentive to keep with it. The important point here is that, if I want to enlist the potential scorn of others in this way, it is not much use simply announcing my opinion of what is best; I need to announce my intention.

(v) Explaining cases of weakness of will without inner conflict. Some people say that in cases of weakness of will there must be an inner conflict: the agent must feel the pull of the course that they are weakly abandoning. And they take this as an argument for the traditional account.[17] That strikes me as quite wrong. Of course, when we self-ascribe weakness of will, we shall typically feel some tension; otherwise, we would not know to self-ascribe it. But we can ascribe weakness of will to a third party in whom there is no conflict. We surely can ascribe weakness of will to a person who has vowed to give up smoking, and who blithely starts up again straight away, saying that they have changed their mind (again, we would need to fill in the details in the right way). We can make perfect sense of this on the view proposed here: we simply say that they gave up their intention too easily. We need to be sure that the initial intention really was formed, and too sudden a change might make us doubt that; and the revision itself might be seen as involving a kind of conflict. But there need be no internal conflict at all at the point at which they smoke the cigarette.

Thus it seems that rather than be a point in the traditional account's favor, its requirement that there be internal conflict actually counts against it. For it is hard to see that the traditional account could count our recidivist smoker as weak willed unless there was, deep down and repressed, a belief that smoking was the wrong thing

[17] For instance, Christopher Cordner, "Jackson on Weakness of Will," *Mind*, XCIV (1985): 273-80.

to do. Yet we do not need to attribute such self-deception in order to attribute weakness of will.

Note that I have not said that cases of *akrasia* are impossible. Suppose that I were tempted by the arguments of Plato and Aristotle and their latter-day followers, arguments to the conclusion that a person cannot choose other than that which they judge best. Then I might try to exploit the account given here to say that what appears to be *akrasia* is really just an unreasonable revision of intention. But I am not tempted by those arguments. So I can happily accept that there are cases of *akrasia*, and that they do give rise to internal conflict. I simply claim that unless they also involve an unreasonable revision of intention, they are not cases of weakness of will.

(vi) Accommodating cases of both akrasia *without weakness of will, and weakness of will without* akrasia. So far we have considered cases of weakness of will without *akrasia*, and we have considered cases of *akrasia* without weakness of will. There are some cases that display both. Consider the following three examples.

(a) Ravi has devoted his life to his poetry and still considers it the most important thing in the world. But he has fallen in love with Una, an English schoolgirl. She finds herself pregnant, and, fearing that her father will force her into an abortion, they elope. Ravi vows that no one will harm the child. His commitment is short lived, however. They are found, and the police threaten him with prison. Faced with the choice between standing by Una and the unborn child, or leaving her to a forced abortion and going free to pursue his poetry once again, he shamefacedly chooses the latter.[18]

(b) Christabel, an unmarried Victorian lady, has decided to embark on an affair that she knows will be disastrous. It will ruin her reputation, and quite probably leave her pregnant. Moreover, she considers it morally wrong. So she thinks it not the best option on either moral or prudential grounds. Nevertheless, she has resolved to go ahead with it. At the very last moment, however, she pulls out: not because of a rational reconsideration of the pros and cons, but because she simply loses her nerve.[19]

(c) The President has his finger on the nuclear button and his threat is simple: if the enemy invades, he will retaliate in massive

[18] Rumer Godden, *The Peacock Spring* (London: Macmillan, 1975).

[19] Thanks to Rae Langton for this example. It is adapted from the story of Christabel LaMotte in A.S. Byatt's novel *Possession* (London: Chatto and Windus, 1990)—Byatt's Christabel, however, does *not* lose her nerve. In thinking about the case, it is important to recall the distinction, made above in discussing the hang-gliding example, between a revision that it is rational to make, and a revision that is made rationally.

fashion. That will be a catastrophe for everyone, for the President and his people as well as for the enemy. But he reasons that such a threat is needed to deter; and that in order for the threat to be credible it must be genuine: he must really form the conditional intention to retaliate if the enemy invades. Then the unthinkable happens: the enemy invades. Enraged, the President prepares to retaliate. But then, faced with doing what he takes to be morally and prudentially wrong, he hesitates, revises his intention, and desists.[20]

All three examples have a similar form. At some point, each agent either forms, or finds themselves with, an intention that is in some sense *akratic*: an intention to do other than what they judge best. Ravi intends to stand by Una; Christabel intends to have the affair; the President intends to push the button. None at this point seems to show weakness of will, notwithstanding the gap between what they intend and what they judge best. So here we have *akrasia* without weakness of will; or at least, we would have if the agents had acted on those intentions. Then the agents revise their intentions, bringing them into line with what they judge best. Ravi now intends to abandon Una and pursue his poetry; Christabel now intends to forego the affair; the President now intends to take his finger off the button without pushing it. But in each case, it is precisely the revision that leaves them open to the charge of weakness of will. So here we have weakness of will without *akrasia*.

For Ravi and Christabel, I think it is clear that revising the intention constitutes weakness of will. There are, however, important differences between the two examples. Ravi has a guilty conscience about what he is doing; while he does judge it the best option, all things considered, he does not judge it to be the morally best option. So might we try to save the traditional account by arguing that weakness of will consists in failing to do what one judges to be morally best? Christabel's case shows us that this defense will not work. She does think that the best option, all things considered, is the morally best option, yet in revising her intention she, too, shows weakness of will.[21]

[20] I am grateful to David Lewis for this example. I am not sure that he intended it for quite the use to which I have put it.

[21] Advocates of the traditional account have typically insisted that the judgment of what is best, all things considered, is not the same as the judgment of what is morally best; and that the account of weakness of will should be framed in terms of the former notion, not the latter. See Donald Davidson, "How Is Weakness of the Will Possible?" in his *Essays on Actions and Events* (New York: Oxford, 1980), pp. 21-42, here p. 30.

The case of the President is more complicated. It is not clear whether we should ascribe weakness of will to him; but that is because it is unclear whether or not it is reasonable for him to have a tendency to change his mind in this case. Let us return to the point at which the President was still trying to make the deterrence policy work. Then, at least according to the logic of deterrence, it was reasonable for him to form the intention to retaliate if the invasion happened. At that point, we could surely have expressed a conviction that he would stick with this intention by describing him as strong willed; and, equally, we could have expressed a doubt that he would stick with it by describing him as weak willed. That is what the account predicts. But once things have gone wrong, it is not obvious what he should reasonably do, even if we accept the logic of deterrence. On the one hand, he reasonably judged that he should retaliate in case of invasion, and he is now in such a case; on the other, triggering an all-out nuclear war is madness. If we incline to the second of these considerations (as we surely should) then we should not describe him as weak willed if he revises his intention. But the memory of what we said before the invasion should make us uneasy with that judgment. In cases like this, we get ourselves into a very real and very dangerous muddle over what it is reasonable to do; our confusion over applying the notion of weakness of will reflects that muddle.

IV. RESPONSES TO OBJECTIONS

I have given my reasons in favor of the account that I propose. I shall finish by responding to five possible objections.

(i) Why be so strict? Amelie Rorty, in her enlightening paper "Where does the Akratic Break Take Place?" answers her own question by providing a multitude of candidates (*op. cit.*).[22] I have suggested that weakness of will consists in just one kind of failing. But why should I be so strict? The English language is a plastic instrument. It seems overwhelmingly likely that sometimes people have used the expression 'weakness of will' to describe other failings, such as some of the others that Rorty mentions. In particular, is it not very likely that sometimes it has been used to describe the case in which a person fails to form an intention to do what he knows to be best, and as a result does what he knows not to be. In short, is it not very likely that the traditional account captures one of our uses of the expression?

[22] Note that Rorty is here giving a discussion of *akrasia*, not of weakness of will; like me, she denies that the two are the same. See note on p. 333.

Perhaps that is right. If so, I should rest content with the claim that many cases of weakness of will are captured by the account proposed here; I should offer it as a supplement to the traditional account, not as a replacement. But I cannot help thinking that the traditional account is not simply inadequate, but straight-out wrong. First, we have seen a number of cases in which people appear to be *akratic* without being weak willed. Second, I doubt that there are any clear cases of weakness of will that can be captured by the traditional account and not by mine.

Consider this example concerning Jim Dixon, hero of Kingsley Amis's novel *Lucky Jim*.[23] Dixon, a junior and untenured lecturer, is staying with Professor Welch, his head of department and arbiter of his future in the university. He awakens after a night of heavy drinking to discover cigarette burns through his bedclothes.

> Had he done all this himself? Or had a wayfarer, a burglar, camped out in his room? Or was he the victim of some Horla fond of tobacco? He thought that on the whole he must have done it himself, and wished he hadn't. Surely this would mean the loss of his job, especially if he failed to go to Mrs. Welch and confess what he'd done, and he knew already that he wouldn't be able to do that (*ibid.*, p. 63).

Would we not say that Dixon displays weakness of will here? Yet he never forms an intention to tell Mrs. Welch. So must we not accept that there can be weakness of will that does not consist of a failure to do what one intends? This seems like good evidence that the traditional account is sometimes correct. But I am not so sure. Why does Dixon not form the intention to tell Mrs. Welch? Presumably, because he knows that he would not be able to go through with it, if he did. So he knows that, if he did form such an intention, he would display weakness of will, where this is understood in terms of the present account. It is because he knows that he is someone with a tendency to weakness of will that he acts as he does. So, on the account given here, his weakness of will explains his action (or rather his inaction). It seems to me that is good enough. Once we have said that we feel no compelling need to insist that Dixon actually exhibits weakness of will here.

If we are to get a really compelling counterexample, we need a clear case of weakness of will in which a person knows that, if he were to form the intention to do what he judges best, he would stick to that intention; but he fails to form the intention. Perhaps I am blinded by partiality to my account, but I am unable to think of one.

[23] London: Gollancz, 1953.

(ii) Can the account accommodate disagreement? I said that one plausible way of filling out the story of the young Schelling and his blankets is to say that he displayed weakness of will twice over. But there are other things we might say. Knowing Schelling's subsequent career as an economist rather than an explorer, we might say that the real Schelling (the rational realistic Schelling) is the one who wanted to keep the blanket on; hence it was only the fanciful dreamer who displayed weakness of will when trying to get it off. (Contrast Francis Chichester,[24] who, true to his real self, stripped his childhood bed down to just a sheet, seems never to have been tempted to pile on extra blankets, and became a fearless adventurer.) Alternatively, we might say that the weakness of will occurred only in the middle of the night; the real Schelling was the one who made his decisions in the light of day, not tempted by the lure of the warm blanket. We can imagine that debate about this could go on for some time. We can imagine the kinds of evidence that would be relevant. And we can even imagine that, when all the available evidence was in, people might still disagree.

It is easy to see how the traditional account of weakness of will can try to explain this disagreement. The debate would be over what Schelling really valued: his comfortable sleep or his training as an explorer. Can the account offered here explain it as well? I think it can; indeed, I think it can actually do better than the traditional account. The account offered here employs both a descriptive and a normative element. To display weakness of will, an agent must have formed an intention that he revises; that is the descriptive element. And his revision must have been something that, by the standards of a good intender, he should not have done; that is the normative element. Disagreements about whether or when agents manifest weakness of will typically result from disagreements about the second of these two factors. So we might argue that Schelling should have kept to intentions that were formed when he could think clearly (such as in the daytime) rather than when he could not (such as when he was tired at night). Or we might argue that he should have kept to intentions that were realistic (the intention to get a good night's sleep) rather than romantic and unrealistic (the intention to become an Antarctic explorer).

Here, we have the second way in which the vagueness in the account mirrors the vagueness in the concept. I doubt that there will, in general, be agreement on when an agent shows weakness of will; and I suspect that this disagreement is actually better captured by

[24] *The Lonely Sea and the Sky* (London: Hodder and Stoughton, 1964), p. 26.

the account on offer here than by the traditional account. On the traditional account, the answers will depend on what the agent thought best, and that is a purely descriptive question. On the account on offer here, the answers will depend on which intentions he should have stuck with as a rational intender. That is a normative question. I suspect that the deep-rootedness of the debate gives us some reason for characterizing it as normative, but I cannot think of a way of resolving the issue with any degree of certainty.

(iii) Does the account work for policy intentions? It might be thought that the account cannot accommodate weakness of will with respect to policy intentions. Suppose I intend to get up every morning at six. Then can I not show weakness of will by lounging in bed today till nine, even though I have not revised my policy intention? The right answer is that I have revised the intention. I have not abandoned it altogether, but I have revised it. I have inserted a let-out clause: I intend to get up every day at six, *except today.*

(iv) Can the account accommodate actions performed without intentions? I have written as though there is a tight link between intentions and actions: as though we do not get actions without intentions. Perhaps that connection is too tight. Suppose it is. Then it seems that the following is possible. I form an intention to perform some act, and then do something else. But I do not revise my intention; I simply act without an intention. Could that not be a case of weakness of will that does not fit the account offered here?

If there were such cases, I doubt that they would pose too much of a problem; we could extend the account to embrace certain failures to act on intentions as well as revisions of intentions. (Not all such failures should be included: I do not typically display weakness of will if I *forget* my intention.) But I am not convinced that there are any cases of this kind. Let us consider one of Bratman's examples of action without intention. Suppose someone throws me a ball and, without thinking, I just reach out and catch it. That is perhaps an intentional act; but it is not clear that I formed an intention to catch it. I acted quite without reflection.[25] But now suppose I form the intention not to catch the ball; you throw it, and again I just reach out

[25] Bratman, *Intention, Plans, and Practical Reason*, p. 126. He also discusses actions that do not result from an intention but from what he calls an "endeavoring" (chapter 9). If we follow him in thinking that the distinction is necessary, shall we need to augment the present account by substituting 'intention or endeavoring' in place of each occurrence of 'intention'? I am not sure, but I think not, since, on Bratman's account, each endeavoring to do something which is not an intention to do it will nonetheless be a intention to do some other thing (p. 119); so each over hasty revision of an endeavoring will also be an over hasty revision of an intention.

and catch it. Have I shown weakness of will? I rather doubt that I have. I suspect that in so far as the act really is one that is done without an intention, it is one that falls outside my immediate control, and so outside of the area in which weakness of will can be shown. I might resolve not to faint when the incision is made, not to duck when the jets fly low overhead, not to smile when my friend inadvertently makes a clown of himself. If I cannot help myself do these things, however, my failure to stick by my resolution does not manifest weakness of will. Certainly, these are failures of self-control. But this is just another reason for not equating such failure with weakness of will.

(v) Can the account distinguish weakness of will from compulsion? It is a commonplace in the philosophical literature that weakness of will should be distinguished from compulsion. On the traditional account, the distinction is made as follows: one displays weakness of will when one *freely* acts against one's own best judgment; one is compelled when one's action against one's own best judgment is *not free*.[26] But it might seem that the account offered here can make no such distinction. Consider a compulsive person—a kleptomaniac, for instance—who intends never to steal again, but who revises that intention and does. Will his revision not be one that he rationally should not have made? Yet if so, he will count as weak willed.[27]

I have two things to say about this. First, it is not obvious to me that it would be a disaster if the account did classify compulsive acts as weak willed; for it is not obvious to me that they are not. Certainly, we would not normally *say* that a compulsive person was weak willed, but that could be because it would be inappropriate rather than false to say so—in the way that it would be inappropriate to say that sadistic torturers are unkind. Are we really averse to saying that compulsives are *pathologically* weak willed?

Second, it is not obvious whether the account offered here would classify compulsives as weak willed. Certainly, one can imagine cases for whom the compulsive action is so automatic that there is no need of an intention, and hence no need to revise any contrary intention: for the kleptomaniac who simply finds himself placing objects in his

[26] See Watson; Kennett and Michael Smith, "Philosophy and Commonsense: The Case of Weakness of Will," in Michaelis Michael and John O'Leary-Hawthorn, eds., *Philosophy in Mind* (Dordrecht: Kluwer, 1994), pp. 141-57. Watson argues that the relevant notion of a free action is one which the agent could have performed had they developed in themselves the kinds of skills and capacities that we expect them to develop.

[27] Thanks to Smith and Linda Barclay for making me think about this issue.

pocket, intentions to steal or not to steal are quite beside the point. So such cases will not be classed as weak willed. But what of the kleptomaniac who consciously but compulsively does revise his intention not to steal; that is, who is simply unable to refrain from making the revision? Will he count as someone who has an unreasonable tendency to revise, and hence as someone who is weak willed? This depends, of course, on how we determine what is reasonable. We could use internal criteria, criteria that always restrict what it is reasonable for an agent to do to that which it is in the agent's power to do.[28] Such criteria would not classify the kleptomaniac's revision as unreasonable, and hence would not classify him as weak willed. Alternatively, we could use external criteria, criteria that do not restrict what it is reasonable for an agent to do to that which it is in the agent's power to do. Such criteria probably would class the kleptomaniac's revision as unreasonable, and hence probably would classify him as weak willed. I have not come down in favor of one or the other of these criteria of what is reasonable; indeed, I suspect that while I am trying to give an account of our ordinary notion of weakness of will, I should remain uncommitted. For I suspect that this is another place where our ordinary concept is vague.

So the present account of weakness of will has difficulties with compulsion only if two further claims are met: if it is right to use external criteria of what is reasonable, *and* if the compulsive should certainly not count as weak willed. I have questioned both of these claims; but even if they were true, a simple amendment to the account would fix things up. Rather than say that weakness of will consists in overreadily revising an intention, say that it consists in overreadily revising an intention *when it is in the agent's power to desist from that revision.*

RICHARD HOLTON

University of Sheffield

[28] In the agent's power in some broad sense; we could follow Watson in thinking that what is important is what would have been in the agent's power if the agent had developed the capacities that we normally expect them to develop.