

BOOK REVIEWS

Notes

¹ In the preface, Ruben explains, “I do not think that use of grammatical gender in personal pronouns or possessive adjectives reveals anything of any importance, and so I use ‘his’, ‘he’, and so on, to indicate persons of any, or no, gender in the sexual sense. I hope that the reader sensitive to these issues will bear that in mind” (Ruben 2003, viii). As it makes no difference to the author, I am helping myself to ‘her’ and ‘she’ throughout the review.

² Donald Davidson, “Agency,” in *Essays on Actions and Events* (1971; repr., Oxford and New York: Clarendon Press, 1980), 43–61, 59.

³ Davidson, “Agency,” 59.

The Philosophical Review, Vol. 114, No. 1 (January 2005)

Sarah Stroud and Christine Tappolet, eds. *Weakness of Will and Practical Irrationality*. Oxford: Oxford University Press, 2003. Pp. x, 317.

This valuable collection is devoted to a wide range of problems in moral psychology and the philosophy of practical reason, connected primarily with weakness of will. The quality of the papers is consistently high, and they hang together well. The first eight might be the basis for an excellent graduate seminar, using the theme of weakness to investigate such topics as freedom (in the paper by Michael Smith), strength of will (Richard Holton), collective agency (Philip Pettit), emotion (Christine Tappolet), practical judgment (Sarah Stroud), desire (Sergio Tenenbaum), decision and cognitive agency (Gary Watson), and practical rationality (Ralph Wedgwood). The book is filled out with three papers that do not address the issue of weakness: Duncan MacIntosh on prudence, Joseph Heath on decision theory, and Ronald de Sousa on “emotional irrationality.” Their topics fall under the general rubric of “practical irrationality,” and while the collection would be more unified without them—it might simply have been called “Weakness of Will”—they are all worth reading, so that it is hard to complain about their inclusion here.

Most of the papers in this collection treat weakness as *akrasia*: free intentional action against one’s better judgment. Holton is an exception, in his compelling investigation of the empirical and philosophical significance of “will-power.” Building on his earlier work, he makes a useful distinction between *akrasia* as acting against one’s better judgment, and weakness as “over-readiness to abandon one’s resolutions” (41). I am akratic but not weak-willed if I stick to a resolution that conflicts with my better judgment; and I am weak-willed but not akratic if I am unable to keep that resolution, no matter how I try. Holton is interested in *strength of will*, insofar as it corresponds to weakness, rather than *akrasia* (though, of course, the two may coincide). His argument is that, in order to accommodate strength of will, we have to recognize the exertion of the will as a distinctive component of agency, and thus we must revise the belief-desire, and belief-desire-intention models in moral psychology. (He

BOOK REVIEWS

calls these, respectively, “the Humean account” and “the Augmented Humean account.”) Holton’s principal objection is that they mistake the phenomenology of strength of will, conceiving it simply as the victory of one desire (or intention) over another (47–49); what about the *effort* involved in sticking to one’s resolutions? What follows is an illuminating synopsis of empirical research on will-power as a kind of concentration, a mental activity that requires energy and generates such phenomena as exhaustion of the will or “ego-depletion” (56).

I am enthusiastic about Holton’s naturalistic approach, but in the end I did not see the need for revision of the Augmented Humean view. What his argument seems to show is that we need to recognize a certain range of mental activities that make it more likely that one will act on a settled intention—such things as distracting oneself from tempting thoughts, reminding oneself of the importance of one’s task, and so on. As Holton argues, these activities cannot be reductively explained in terms of the strength of one’s intentions or desires (55, 60); but in this respect, they are no different from other mental activities like imagining something, or trying to remember a tune. Nor can they be explained in terms of a strong desire to act on one’s resolution (60–61). They are activities one goes in for *while* one intends to do something else, and *as* one is doing it, which make it more likely that one will finish what one set out to do. But why should acknowledging new action-types affect our theory of the explanation of action in general? Why is there a threat here to the Augmented Humean view? There would be a problem only if the exercise of will-power affects how likely it is that a desire will overcome one’s resolution, without affecting the *strength* of that desire or that resolution. But Holton does not provide a compelling argument for this. In its absence, our picture may be one in which, having formed a resolution, there are various mental actions one can then perform, in order to control the strength of the temptations that might prevent one from acting on it. Not everyone thinks to perform these actions, and some are better at them than others. Nevertheless, allowing for the relevance of beliefs, as well as intentions and desires, the explanation of action still accords with the principle “that it is the relative strength of the conative inputs that determines what the agent will do” (40).

This is not to minimize the importance of Holton’s paper, or of his other work on this topic. Indeed, I sometimes wished the other contributors would pay more explicit attention to the (apparent) distinction between weakness and *akrasia* with which Holton begins. Michael Smith’s engaging paper on “rational capacities” not only blurs that distinction, it also says surprisingly little about the mechanisms of weakness and strength of will. It is concerned with a more abstract question, made pressing by Gary Watson’s classic essay, “Skepticism about Weakness of Will”¹: how should we interpret the claim that someone “could have resisted” a certain desire, which accounts for the difference between weakness and compulsion? Intuitively, the compulsive agent *cannot*

BOOK REVIEWS

resist the desire on which she acts, while the akratic agent *can*—but chooses not to. This contrast matters to their respective culpability.

Smith's strategy for explaining the distinction, and the relevant sense of "could," appeals to a possible worlds comparison between the akratic agent and the compulsive. (I set aside his parallel treatment of recklessness.) According to the basic proposal, which is steadily revised, the akratic agent could have resisted her desire, while the compulsive agent could not, because a possible world in which the akratic agent resists is closer to the actual world than any possible world in which the compulsive agent resists (cf. 22, adapted to the present case). Smith makes two refinements. First, in order to focus on the capacities of the particular agents, and to avoid problems analogous to those about "finkish dispositions," we should abstract away from anything that could have an effect on what our agents do, except their own intrinsic properties (22–26). Second, in order to avoid problems about flukes—for example, the case in which it would take only a small microphysical change to remove the compulsive agent's desire—we should compare not only possible worlds in which agents face the actual choice with which we are concerned, but ones in which they face similar choices, and we should require that the relevant counterfactuals be explained by a "common structure" (26–28).

This is an ingenious solution to a neglected problem. Two features of it seem problematic. First, as I understand it, the proposal has a comparative form. This is clear in its initial presentation, and while it is less obvious in later versions, I think it must be true of them, too. Consider, for instance, Smith's final account of the distinction between weakness and compulsion:

We begin by abstracting away from all those properties that could have an effect on what the weak woman and the compulsive woman desire except the relevant properties of their brains. We then note that the weak woman would, whereas the compulsive woman would not, have a host of counterfactuals true of her. She would desire to refrain from similar drinks, and the like ... in those nearby possible worlds in which she believes [she should do so]. (35)

What does Smith mean by "nearby possible worlds"? The claim cannot be that the weak-willed drinker refrains from having another drink in the *closest* possible world in which she thinks she should do so, for that is the actual world, in which she does not refrain. Nor is it plausible to say that she refrains from taking another drink in the closest *similar* worlds, for it may be that she would take a drink, if offered one, under *most* circumstances. The point has to be, as in Smith's initial claim, that the worlds in which she decides not to take a drink are *closer* than those in which the compulsive agent does so. The difficulty, then, is to map this comparative claim—in effect, that it is "more possible" for the weak-willed agent to act as she should than it is for the compulsive agent—on to noncomparative claims about capacities. If we had an absolute standard for "sufficiently close worlds," we could simply look at whether the agent resists her desire in worlds that meet that standard, concluding that she is weak-willed if

BOOK REVIEWS

she does, and compulsive otherwise. But where does the standard come from? It was part of Watson's point in "Skepticism about Weakness of Will" to suggest that standards of this kind can be found only in what is normal hereabouts. Watson concludes that they are socially relative—they may be different in Athens than in Sparta. Smith does not address the question of noncomparative standards, and so he leaves the suspicion of relativism in place.

The second problematic feature of his account is that it seems to ignore the possibility of the utterly determined or resolute akratic. This is one place in which Holton's distinction is useful. If we conflate the judgment of what one should do with the intention to act, it may seem a rather fragile matter that the akratic agent acts as she does. She must be conflicted, and it would not take much for her to mend her ways. But once we distinguish these things, we make room for the unconflicted akratic, who has no intention whatever of doing what she knows she should do. The worlds in which she would do so, in this sort of circumstance, are very distant indeed—as distant as those in which the compulsive agent is not compelled.

It might be said, in Smith's defense, that the distinction between weakness and *akrasia* is not as simple as this, or that, even if we make this distinction, it is a necessary truth that agents tend to act on their all-things-considered judgments. I'm not sure that the latter point would undermine the objection made in the previous paragraph. But it leads me to what is perhaps the central theme of the present volume: a sympathy for what Tenenbaum nicely calls "scholastic" conceptions of agency and desire (148). We can distinguish two "scholastic" claims about motivation: *internalism*, according to which it is constitutive of rational agency (the capacity to act for reasons) that one have some tendency to act on, or to be moved by, one's evaluative judgments; and the claim that intentional action (or, perhaps, desire) must take place "under the guise of the good."

Stroud's paper is a sustained argument for internalism about "practical judgment" (122), which she identifies, cautiously, with the judgment of what one has most reason to do. (More precisely, Stroud argues that there is such a thing as practical judgment, where something counts as practical judgment only if internalism is true of it.) Her main argument for a necessary connection between judgment and motivation (or the tendency to act) is that, if the connection is purely contingent for an individual agent, the state of affairs in which *no one* has any tendency to be moved by their practical judgment must be "coherent, intelligible, and conceptually possible" (143). But this state of affairs is *not* possible; so, internalism is true (143–44). In effect, her argument appeals to a principle of recombination for possibility: if it is possible for one agent to be wholly unmoved by practical judgments, it is possible for everyone to be that way. While I agree with Stroud about the dubious coherence of the global case, I think the inference here is more problematic than she makes it seem. Consider the following view: in the paradigm case, it is a condition of the capacity for practical judgment that one have some tendency to act on the judg-

BOOK REVIEWS

ments one makes; however, it is possible for wholly indifferent agents to grasp the relevant concepts, in a way that is parasitic on the paradigm case, much as the blind person's possession of color concept is (often thought to be) parasitic on that of people with sight, or the layman's use of medical concepts is (arguably) parasitic on that of medical experts. If this makes sense, we can accept a kind of weakened internalism, while allowing for the possibility of thoroughgoing (but perhaps not global) indifference to reasons.

In any case, I think the more significant doctrine here is not internalism but "the guise of the good." Versions of this idea are defended by Tappolet, Tenenbaum, and Watson. But what they say on its behalf seems rather thin. For instance, Tappolet claims that it is "simply not clear that an agent can aim at a value that she considers to be a bad or insufficient reason for performing [an] action" (118). This is *very* strong indeed, and Tappolet defends it only by giving a controversial example. Tenenbaum makes a weaker, but still contentious claim:

In a proper intentional explanation, the agent (or a third person) will be able to explain the point of engaging in such an activity; in other words, he will be able to explain what good he sees in the pursuit of this activity. (150)

The transition from the first clause of this sentence to the second—from having a point to being seen as good—is not self-evident; it needs to be argued for. A final example: in his paper on the will, Watson assumes that practical deliberation is "*reasoning about what is best (or satisfactory) to do with a view to making up one's mind about what to do*" (175). Alert to the possibility that the akratic agent may deliberate, too, Watson urges that even the akratic agent's deliberation may be sensitive to evaluative judgments other than the one he is explicitly acting against (183–84). This is an important point. But first, the fact that it *may* be sensitive in this way does not show that it *must* be. And second, the real issue is more general: whether evaluative judgments (as opposed to mere intentions or desires) must figure in the practical reasoning even of quite rational, non-akratic agents. This collection would have benefited from a contribution by someone like Harry Frankfurt or Michael Stocker, to resist the assumption that agency must be understood in evaluative terms.

I have been focusing primarily on areas of disagreement. Let me say again that this book is a valuable one, full of stimulating arguments and imaginative proposals. Some of the papers here should count as essential reading on their topic. Those in the field will profit from all of them; those outside it will gain a vivid sense of current work at the intersection of ethics and the philosophy of mind.

KIERAN SETIYA

University of Pittsburgh

Notes

¹ Gary Watson, "Skepticism about Weakness of Will," *Philosophical Review* 86 (1977): 316–39.