For An Amoral, Dispositional Account of Weakness of Will

STERLING HARWOOD
San Jose State University

§1 Introduction

The topic of weakness of will (WOW) has attracted a significant amount of philosophical analysis within the last decade or so.1 But a standard approach is still to analyze WOW as basically some kind of action contrary to one's better judgment (ACBJ, for short). I shall argue that WOW need not involve, and often is not, ACBJ, and thus should not be defined as such conduct. My argument is twofold. First, in §2, I will argue for WOW without conduct (i.e., merely dispositional WOW). Second, in §3, I will argue against the normative standards in Robert Audi's conception of WOW, and for an amoral conception of WOW. Further, I shall argue that WOW is a particular kind of disposition to fail to act in accord with one's highest order desire (e.g., one's second order desires about first order desires such as hunger and thirst). Our topic, after all, is weakness of will. So I want to avoid assuming, as ACBJ does, that WOW fundamentally concerns judgment, which is identified with the intellect as distinct from the will. I want to avoid assuming the conflation of conative states (which involve desires, dispositions and the will) with cognitive states (which concern beliefs, judgments, and the intellect).

§2 Dispositional WOW

In this section I shall argue for merely dispositional WOW. If WOW can be merely dispositional, WOW need not involve action, and so a fortiori WOW need not involve action contrary to one's better judgment (ACBJ, for short). ACBJ is the definition of WOW which many (e.g., Donald Davidson) defend, and which Audi and I attack (though Audi uses ACBJ as a part of his definition of WOW).

ACBJ is misleading from the start. ACBJ is a misnomer for WOW, since "better judgment" suggests there must be two judgments, one better than the other. But WOW does not require more than one judgment (even

counting all prior judgments about whatever option is in question). Audi recognizes this when he says:

Whether one's will is weak at a given time should be determined in relation to the practical judgments one holds at the time. It does not depend on which of one's judgments about a given option is better, nor does it ever depend simply on what decisions one has made (and not abandoned) prior to the time in question.²

Audi's quoted view follows his general strategy of developing a conception of WOW that preserves the distinction between "vacillation of intellect" and WOW.³ I agree with Audi that WOW essentially involves something other than merely changing one's beliefs. All I have read who advocate ACBJ agree with Audi here. So we should beware of interpreting ACBJ as classifying the following case as one of WOW. Al wholeheartedly acts in accordance with his present evaluative judgment about his options. But Al changed his judgment prior to his act and his prior judgment was better than is later judgment with which his acts accorded. ACBJ does not say Al's act was WOW, regardless of how his previous judgment was better (e.g., better accuracy, better information, better deliberation, better morally, or better prudentially).

The other distinctions Audi wants his conception of WOW to preserve are the distinctions between: 1) WOW and irresolution (he says irresolution does not imply WOW); 2) acting contrary to one's better judgment and acting inconsistently with one's better judgment (only the former requires awareness of one's judgment when acting); 3) WOW and imprudence; and 4) WOW and compulsion (or coerced action), since he says an unfree will cannot be weak.⁴ I cannot discuss all these distinctions here, but it may help interpretation of our views to note them and note that I fully agree with only 2). In §3 I discuss 4). Finally, note that ACBJ means WOW is action contrary to one's present judgment, of which one is aware, as to what option is better than all other options.

Audi and I reject ACBJ and advocate a merely dispositional conception of WOW. Audi argues:

---

³ Audi, p. 178.
⁴ Ibid., pp. 176f.
Suppose S judges that he should not take a drink and quite consciously tried to resist doing so. He may still form the intention to take one. May he not have thereby exhibited weakness of will? Even if he is not able to take one because the bottle is empty, he has already failed in the kind of inner struggle that often precedes incontinent action (i.e., action manifesting WOW). Similarly, if S makes the judgment at $t$ that it would be better to do $A$ than anything incompatible with doing it, then provided he believes he can do it, his failure to form, at (approximately) $t$, an intention to do $A$, would manifest weakness of will . . . the same kind of failure of will.\textsuperscript{5}

I advocate Audi's description here of WOW as failure of will, an inefficacy of will. The failure can be manifested in inaction and hence ACBJ is inaccurate. One may object that S's forming the intention to take the drink is an act, and so Audi's example comports with ACBJ. I grant that there are mental acts (e.g., mathematical calculations) and that one can form an intention through the mental act of deliberation. But reducing ACBJ to such acts is still a major refinement of WOW which Audi achieves, having shown that WOW need not involve overt acts. Moreover, Audi's example of manifesting WOW by failing to form an intention shows that inaction can manifest WOW. And one cannot rebut the point by saying that the so-called inaction was really the action of refraining or deciding not to act overtly. For the inaction was, by hypothesis, due to not forming an intention rather than forming an intention to refrain from acting. And the failure did not involve deciding to refrain because, by hypothesis, S makes the judgment (i.e., decision) that it would be better to do $A$ (which is the act S failed to form the intention to do).

And I would take Audi's analysis further. One might not completely fail to form the intention to do $A$; one may form it but only after too much of a delay or too much of what Audi calls an inner struggle. "Too much of a struggle" means that even though the result of the struggle conforms to S's better judgment, the struggle was more agonizing, involved stronger desires (on the losing side of the conflict) than those that would have been involved had there been no WOW (i.e., than if one had followed one's highest order desire, as we shall see). This is what Audi means by the kind of inner struggle that normally precedes incontinent action. The struggle (which manifests WOW) can occur even though incontinent action does not. In other words, in WOW the struggle is not always won by the desires conflicting with one's better judgment (or higher order desire, as we shall see).

\textsuperscript{5}Ibid., p. 181. (emphasis added, except for 'failure'.)
Audi’s view, which I endorse on these points, does not say that acting without WOW requires wholehearted action. Only a certain kind of motivational struggle constitutes WOW. Not just any struggle constitutes WOW. He says:

"It is not reasonable to regard as incontinent just any want of S's that is at odds with his practical judgment. S's want to do A may be too weak to be at all likely to bring him to do or even intend to do it; hence, even if doing A is against S's practical judgment, his wanting to do it does not imply that his will has failed in its function (which is making one's second order desires effective (i.e., cause action aimed at fulfilling the second-order desire)). But what if the want, though not predominant, is "strong"... We do not and should not assume that one ought not to want even strongly, things whose realization is against one's practical judgment. For... we regard it as quite natural to want, even strongly, things which we judge we ought not to have, e.g., sweets."

So Audi does not require wholeheartedness to avoid WOW. The strong-willed person can still be tempted to eat sweets. And at least in more serious matters one can even face an agonizing choice. Given some sets of desires, some choices just are agonizing. WOW occurs only when some desires have more strength or influence over you than they would if you were prudent. For the desires competing against them are weaker than they would be if one fulfilled one's second-order desire (i.e., one's desire..."

---

6 *Ibid.*, p. 183. As we will see, my view leads me to say "that the strong-willed person does not want" instead of "one ought not to want."

7 Quantifying the motivational strength of desires may initially seem somewhat problematic, but commonsense clearly accepts ordinal rankings of preferences. Indeed, the concept of preferring one thing to another implies an ordinal ranking. So we need not try to develop a cardinal ranking of preferences to continue to use the metaphorical concept of motivational strength (which is a quantifiable concept). One may object that motivational strength is so metaphorical that it is not an empirically testable concept, especially considering the lack of cardinal rankings of preferences. And since psychology is an empirical science, one might conclude that motivational strength is psychologically unhelpful, explaining little or nothing about the psychological phenomenon of WOW. But Santas, citing two psychological studies, shows how motivational strength has empirically testable significance. See Gerasimos Xenophon Santas, *Socrates: Philosophy in Plato's Early Dialogues* (New York: RKP, 1979), pp. 213-14.
about how strong one's first-order desires, e.g., hunger and thirst, are). This weakness is WOW.

In the immediately foregoing quotation Audi introduces the ideas of what we "ought not to want" and what is "natural to want" and what is "reasonable to regard as" a failure. These are normative concepts I wish to avoid using; for I advocate a non-moral conception of WOW. I shall present and criticize Audi's overly normative conception of WOW in the next section. The foregoing arguments show how WOW is often merely dispositional, and so show the error of defining WOW as action contrary to one's better judgment.

Consider a final example. Suppose Al is a guard with a weakness for women in that Al could not bring himself knowingly to shoot an attacking woman, even if he had a second-order desire to shoot her, and believed doing so would be prudent and morally required. Fortunately, no woman ever attacks Al. So his WOW never manifests itself in any struggle, or any action or inaction. Still, Al is weak-willed, since he is disposed to act incontinently under certain circumstances. This advances Audi's analysis a step further, showing that the disposition that is WOW need not even involve a struggle.

§3 Against Audi's Normative Standards for WOW

In this section we reach deeper and more important issues. I now argue against the normative standards in Audi's conception of WOW, and for an amoral conception of WOW. Audi holds a conception of WOW which has moral or normative implications. Audi makes a moral—or at least a normative—argument for his view that WOW never involves succumbing to an irresistible impulse. He says:

Weakness of will is surely a normative phenomenon in at least this sense: if S does A incontinently (i.e., due to WOW) then S is in some way criticizable for doing A . . . whether morally, prudentially, or in some other way(s) . . . Now suppose we add the widely believed premise that if S does A under compulsion, he is not criticizable for doing it. It would follow that incontinent actions are never compelled.

---

8 Audi, p. 183.
9 Ibid., p. 179. Audi actually rejects this formulation of the argument. He rejects the second premise formulated as "if S does A under compulsion, S is not criticizable for doing A." But Audi has reservations about this argument only because he advances a group of counterexamples which he admits includes no cases of WOW. Thus, we can easily save the second premise from Audi's reservation, which is irrelevant to WOW, by
I am willing to grant the widely believed premise, which seems to be a special case of "ought* implies 'can"", though the premise has increasingly come under attack in recent years, especially concerning the insanity defense. An irresistible impulse (or compulsion) cannot be resisted, and so, based on "ought* implies 'can"", I am willing to grant here than one cannot correctly criticize someone for failing to resist such an impulse. Still, I think that it is at least sometimes a motivational failure, and hence that it can correctly be called a weakness of one's will. But, instead of adopting Audi's view, we can simply conclude that WOW is not always correctly criticizable. This seems a much simpler and better solution. And since not all cases of WOW involve compulsion, the simpler solution preserves Audi's distinction between compelled action and WOW, which is the rationale for his argument.

On whether WOW is always criticizable, and whether WOW is a moral concept, I side with Davidson, who sees WOW as a metaphysical concept in the theory of action, independent of ethics. WOW is more useful as a metaphysical concept free of moral baggage and entanglements. For psychologists could then use the concept of WOW without venturing into an area where they may consider them somewhat less qualified, moral philosophy. And an amoral conception of WOW will help moral theory and action theory, since we could then use the concept of WOW without begging moral questions. I shall defend a somewhat normative, but amoral, conception of WOW. Since the norm I use is prudence, and since I deny that all WOW is (even prudentially) criticizable, psychologists and psychiatrists can use the concept of WOW with less danger of imposing values upon a client that are not the client's own.

Further, Thalberg gives an initially plausible argument against a moral conception of WOW, saying:

"He shouldn't have done Y... cannot always [be said] of a person who had better [subjective] reasons to do X instead. Otherwise you would be endorsing the dubious

reformulating the second premise as "if S does A under compulsion and if A is a case of WOW, then S is not criticizable for doing A." Audi admits that the originally formulated argument is valid. My revision of the second premise does not change that. My revision merely makes clearer what the original context was—cases of WOW. And Audi accepts the first premise (i.e., "if S does A incontinently then S is in some way criticizable for doing A") and the conclusion (i.e., "incontinent actions are never compelled") of the argument. On all this, see, Ibid., pp. 179-80.

10 Ibid., p. 180.

principle that everyone including Nero, Genghis Khan, Torquemada and Hitler, ought to do what he judges best . . . The point is that we often hold that some action, which another person believes he ought to perform is wrong or inadvisable in some way. If he fails to perform, then we would be inconsistent if we maintained our disapproval of the action he was contemplating, but condemned him for failing to carry it out. Only when we share the agent's belief that he ought to do X is it logically possible for us [consistently and correctly] to decry his failure [as such].

But I must disagree with Thalberg. For we can consistently and correctly condemn Jim Bakker or Jimmy Swaggart for their hypocrisy even if we scarcely share their beliefs about what they ought to do. Specifically, a libertarian or liberal might claim that Bakker and Swaggart are wrong to condemn prostitution among consenting adults, but that if they are going to condemn prostitution anyway, at least they should not pay prostitutes for sex. One can correctly decry their paying for sex even if one thinks prostitution per se is morally permissible.

But we need an amoral conception of WOW, for in at least some cases of WOW we should even praise the would-be Hitler for his WOW. Here is the first of the two main examples in this paper that I wish my reader to focus on. Suppose young Hans is indoctrinated by Nazis and is months later ordered to kill a group of Jews. Hans believes he ought to kill them, but his humanity "gets the better of him." He has lingering, unarticulated feelings about what he ought to do. Still, he believes he ought to kill them, and that it is highly probable that his Nazi teachers and superiors have been right all along. Hans is not a squeamish person. And he has a second-order desire to kill them according to his orders. But come time to pull the trigger, his humanity and sympathetic feelings for the Jews lead Hans to let them escape, even though he has not changed his mind, and is acutely aware of his motivational failing. Further, suppose Hans is prudent to let them escape because Hans has been trusted to kill them, by himself, in an area bordering enemy lines (and far from other Nazis or Nazi-sympathizers) and to dispose of their bodies completely. We can suppose that firing his gun so close to enemy lines (or taking time to dispose of the bodies completely) is imprudent. I see no sound, normative criticism of Hans' act of letting the Jews escape.

The counterargument that we can always criticize someone for acting against his better judgment just because it is against his better judgment (i.e., just on grounds of internal inconsistency) trivializes the thesis. First,

---

the counterargument begs the question against arguments that sympathy should sometimes take priority over conscience.13 Second, I argued above that WOW need not involve inconsistency between superior and inferior judgments. Third, foolish consistency is the hobgoblin of little minds; not all internal inconsistencies are normatively criticizable, but some are—as the cases of Bakker and Swaggart show. Simply saying that all internal inconsistencies are criticizable as such trivializes the thesis, and begs the question by failing to give any further or independent rationale for the criticism. Moreover, this objection would require fanaticism (Nazi massacre) to avoid criticizable WOW, which is absurd. Hans should have conflicting feelings, given his circumstances. Further, even if there is an undiscovered, normative criticism of Hans, we can tell Hans' case is one of WOW even before discovering any criticism. So it is highly misleading and unnecessary to go out on a limb and say WOW must be normatively criticizable and include that as one of the very defining characteristics of WOW.

One may object that I owe an explanation for how the WOW of Hans is not criticizable even though the WOW of Bakker and Swaggart is criticizable. The distinction between the cases is that Hans is young and (understandably) confused whereas Bakker and Swaggart are older (old enough to know better) and hypocritical.

Perhaps all Audi means by saying WOW is always criticizable is that WOW always involves a case where the agent is morally responsible and could have refrained from doing the incontinent act. That is, WOW is always criticizable because were the incontinent act wrong, the actor could be correctly criticized for doing the act, since the stricture "ought" implies 'can" is met in all case of WOW. I consider this to be too watered down to be Audi's view. But if it is his view, then I differ only on case of compulsion (e.g., irresistible impulses). Audi's argument for distinguishing compulsion from WOW is that WOW is essentially a normative concept.

There is another reason, besides Hans' case, to avoid Audi's normative analysis of WOW. Audi classifies the (voluntary) knowing pursuit of evil as WOW, saying: "S exhibits [WOW] . . . if . . . S forms the intention or a predominant want to do A, where doing A is against his better judgment . . ."14 The knowing pursuit of evil is, I grant, acting against one's better judgment. But here is the second of the two main examples I wish my reader to focus on. Surely an evil person, call him Cecil the devil worshipper, can can even do research on what the morally right thing to do would be, and then, based on this research, deliberately and painstakingly do what is evil. Cecil can be moved by prudential considerations (e.g., he can safely steal undetected) and his desire to act on evil desires and do evil.

---

14 Audi, pp. 180f.
WEAKNESS OF WILL

(as such). Moreover, suppose knowingly doing evil fulfills Cecil's second-order desire to do evil as such. Audi would say Cecil exhibits WOW because, by hypothesis, Cecil's moral judgment is that Cecil ought not do the evil. Because Cecil acts wholeheartedly, without motivational conflict and in accordance with his second-order desire to do things based on evil desires, I claim he does not act incontinently. I grant that Cecil acts against his better judgment, but I claim this is irrelevant given that he acts in accord with his second-order desire (i.e., his desire about his other desires).

One may object that Cecil must have some sort of overall judgment that it is really best to do what is evil, since to worship the devil is to believe that the devil's way (evil) is best. Thus, one may object, Cecil's case fails to fit ACBJ, and thus cannot be a counterexample to ACBJ. But this objection conflates conative and cognitive states. Cecil, as a purely intellectual matter, identifies the evil and the good acts to do. Cecil can even intellectually accept a Kantian view that obeying morality is what is more important overall. Yet Cecil can have absolutely no desire to obey morality. He can have a desire to do evil, as such, instead. If Cecil's case is too strange to see this clearly at first, compare the following more mundane case. Pat acknowledges the soundness of arguments concluding that he should give at least something to charity to relieve hunger. But Pat is unmoved, having no desire to do so. Pat asks "What's in it for me?" Pat is apathetic toward morality. Similarly, Cecil is unmoved to obey morality. He can say "What's in it for me? I desire to do evil as such, and prudence does not go against the sort of evil I have in mind." But in the cases of Pat and Cecil where is the motivational weakness required for WOW? They act exactly as they want, and they act wholeheartedly without any motivational conflict. So I conclude that there is no WOW in their cases.

Those who object to my view here have the burden of showing the motivational weakness in Cecil's wholehearted actions. One may object that the motivational weakness is simply the absence of motivation to do the right thing. But this simple view would prove too much. Since nobody is perfect, we would all be weak of will virtually all the time. I cannot defend a particular morality here. But morality seems demanding enough to make this objection classify too many of us as weak of will too much of the time. Further, the objection fails to see how WOW involves a conflict within the agent's psychology. Recall Audi's view about WOW as an inner struggle. A case of wholehearted yet incontinent action seems incoherent.

One may object that I am taking "better judgment" to mean "judgment as to what is morally best" and that this is not the way it would be taken by anyone inclined to accept the definition of WOW as acting against one's better judgment. The objection clearly fails, for I do not so take "better judgment." I mean it just as Audi, Davidson, and Thalberg mean it, namely
as a judgment as to what one ought to do, *all things considered*.\(^{15}\) The point in Cecil’s case is that this judgment about what is best, all things considered, is a purely intellectual judgment that the moral act is best, and that the judgment involves no motivation to do the moral act.

But the ambiguity in the use of “better judgment” by Audi does create some important ground where my view and Audi’s might be reconciled. Audi might mean by “better judgment” one’s *decision*, all things considered, about what one will do. The judgment here is the occurrcnt, present decision to start willing the act in question (i.e., to issue the relevant volition). It is not a mere prediction about what one is about to do. I agree that if one acted contrary to one’s better judgment (in this sense), then one could act incontinently. And to the degree that one was disposed to act contrary to such *decisions* one would have WOW. One can have conflicting second-order desires A only, and desire to have a set of desires B only, yet know that the two sets are incompatible because the sets contain different elements (desires). But in such cases the decision one acts against must be present at the time of action, since I agree with Audi that WOW is distinct from mere intellectual vacillation. If one failed to form the intent to do the act decided upon, one has WOW. Such a failure would occur if one formed instead an intent to do another act (or no act). All decisions, in the sense used here, are accompanied by first-order and second-order desires to do the act decided upon (or, in the case of conflicting second-order desires, a third-order desire—i.e., a desire about one’s second-order desires, particularly a desire about how the conflict between the second-order desires will be resolved). For one is deciding which of one’s first and second-order desires to fulfill. Harry Frankfurt details the role second-order desires play in self-reflection, self-evaluation, and in the issuance of volitions.\(^{16}\) And I at least roughly follow Frankfurt here.

Consider another case. Audi might mean that in some cases the better judgment is one’s judgment about what act all prudential factors favor. One could be not incontinent yet act against such a judgment based on a first-order desire (e.g., an imprudent impulse) or even based on a second-order desire than on lack certain first-order desires. Since prudence is calculated by considering all desires of the agent, event he first-order desires that one has a second-order desire to end, one’s second-order desire can conflict with one’s prudential judgment. (For simplification, I assume they conflict in a case where the agent knows that moral considerations and other non-prudential considerations favor neither option considered by the agent). Such a conflicting second-order desire can be, for example, the desire to do an act that will thereby make one a different sort of person than one is. But given that one is not that sort of person, prudence dictates refraining from

---


doing an act, A, to fulfill the second-order desire, and to do instead another act, B, which prudently satisfies the vast bulk of one's desires, but which fails to satisfy one's second-order desire. If one had a conflicting second-order desire to do B, then I think knowingly doing A instead would be incontinent. But if the second-order desire to do A were one's only second-order desire (and there were no even higher order desires), then I suspect that knowingly doing A would hardly be incontinent. For, on the contrary, the case seems to involve strength of will, where one becomes what one really wants to be (and has a second-order desire to be) through sheer force of will, even though becoming such a person is imprudent. The very fact that one overcame all the other desires that made doing B prudent shows how strong one's will was. So such cases fit what I conclude must be the best definition of WOW, namely, that WOW is the disposition to fail to act in accord with one's highest order desire. As Robert Roberts argues:

Why do we tend to think a moral achievement greater if more difficult? . . . [M]orally difficult actions display. . . . the virtues of will power. But . . . a deeper basis for our feeling here is that the greater the moral obstacles (that is, contrary inclinations) a person has overcome in doing something, the more his action seems to be his own achievement, his own choice, and thus to reflect credit on him as an agent. It seems to show his action is his in a special way.17

The agent should get metaphysical credit for the strength of this effort.

§4 Conclusion

I conclude that not all WOW is criticizable, since Hans' case is a counterexample to Audi's view. And using Cecil as a counterexample I argued against Audi's classification of the knowing pursuit of evil as WOW. Those who object to my view have the burden of showing 1) what is immoral in Hans' action of letting the Jews escape, and 2) what is motivationally weak in Cecil's wholehearted actions. My main conclusion is that WOW is having a disposition (to fail to act in accord with one's highest order desire) that is greater than such a disposition would be (at least typically) if one were completely prudent. One may object that my view leaves no rational, motivational role for morality. But prudence is at least normative, and thus my conception of WOW, although amoral, is

normative and hence partially reconciled with Audi's view. Moreover, as I said earlier, I mean prudence in the sense of maximizing the satisfaction of all of one's desires, including one's desires about doing the moral thing as such. So I leave room for a rational, motivational role for morality, since I preserve the distinction between prudence and egoism (by including in prudence any desires to do the morally right thing as such). But while my conception of WOW is somewhat normative, it is still an amoral conception of WOW, because I showed how morality is irrelevant to the definition of WOW. My conception of WOW is only somewhat normative, since I conclude, based on my example of Hans, that nothing criticizable or imprudent is a necessary condition of WOW. But I conclude that following the norm of prudence, except in the case at the end of §3, which involves willing to become a new sort of person, is a sufficient condition for strength of will.18

18 I wish to thank the following for their comments: Carl Cinet, Gordon Graham, John Heil, Alfred Mele, William Rottschaefer, Natalie Vania and Burt Warner.