Practical Reasoning and Weakness of the Will

MICHAEL BRATMAN

STANFORD UNIVERSITY

In a case of weak-willed action the agent acts—freely, deliberately, and for a reason—in a way contrary to his best judgment, even though he thinks he could act in accordance with his best judgment.[1] The possibility of such actions has posed one problem in moral philosophy, the exact nature of the problem it poses another. In this essay I offer an answer to the latter problem: an explanation of why a plausible account of free, deliberate and purposive action seems to preclude the possibility of weak-willed action. I then try to resolve the first problem by developing this account in a way which allows for this possibility. The possibility of weak-willed action is made problematic by an account which sees free, deliberate and purposive action as involving the conclusion of a piece of practical reasoning. Solving the problem does not require us to abandon this conception but, rather, to notice certain special features of the relation between premises and conclusion in such reasoning.

I

Weak-willed actions, as I have defined them, do not exhaust the range of phenomena which might be said to be instances of weakness of the will. They do not include cases of what Aristotle called impetuousity, since impetuous conduct contrary to one's best judgment will not be deliberate conduct. Nor do they include cases of non-action which are not deliberate omissions, for example: staying in bed while knowing one ought to get up.

What are included are those cases most directly threatened by certain plausible theses in the theory of action.

Focusing on these cases will help us understand why weakness of the will is problematic.

II

Suppose I am considering whether to play basketball or read my students' term papers. I see considerations in favor of each course of action. On the one hand, playing basketball would be fun and a welcome change of pace at the end of the term. On the other hand, the papers must be read by next week and if I wait much longer I will not have time to give them adequate attention.

I think that playing basketball would be, *prima facie*, the best thing to do, given that it would be both fun and a welcome change of pace. I also think that reading the papers would be, *prima facie*, the best thing to do, given that it would insure that they received adequate attention. Seeing no other relevant considerations, I weigh these conflicting considerations in order to settle the question of what to do. Thinking my professional duties to be very important, I conclude that the latter considerations outweigh the former ones and, so, that it would be best to read the papers. So, with a slight sense of loss, I proceed to read them.

I read the papers deliberately and for a reason. Being under no complusion so to act, I do so freely. This is a case of free, deliberate, purposive action. I shall call such actions *full-blown* actions.

In this case I act on the basis of a piece of practical reasoning which includes a consideration of my reasons for acting as I do. In this respect my case is similar to a case of reaching a new belief on the basis of a piece of theoretical reasoning. Suppose, for example, I notice the dark clouds outside and then hear a weather report which says it will rain today. On the basis of this evidence I conclude that it will rain today. My new belief that it will rain is a conclusion of a piece of reasoning which included a consideration of my reasons for this belief. This suggests that we treat full-blown action as analogous to reaching a new belief on the basis of theoretical reasoning. Full-blown action is the conclusion of a piece of practical reasoning which includes a consideration of (some of) one's reason(s) for so acting.

This may be too simple, however. We may want to say, rather, that it is the intention or, perhaps, the "volition" so to

act which is the conclusion of such reasoning in a case of full-blown action. So let us say only that full-blown action involves the conclusion of an appropriate piece of practical reasoning.

In one sense a piece of reasoning is a psychological process. In another sense, a piece of reasoning is an abstract structure consisting of certain sentences or propositions or, perhaps, certain proposition-like entities.[2] In the first sense, my reasoning to the conclusion that it will rain is a process involving certain beliefs of mine. In the second sense it is an *argument* consisting of the sentential (propositional) contents or expressions of these beliefs. When there is reasoning in the first sense we can construct a corresponding argument, consisting of the contents or expressions of the psychological entities involved in that reasoning. Hereafter, when I speak of reasoning I shall mean reasoning in this first sense. Reasoning in the second sense I shall call an argument.

My state of believing it will rain is the conclusion of my reasoning. In contrast (the proposition expressed by)

It will rain,

is the conclusion of the corresponding argument. The conclusion of my reasoning corresponds to the conclusion of the corresponding argument.

Similarly, my full-blown action, or my intention or volition so to act, will correspond to the conclusion of the argument which corresponds to my practical reasoning. This conclusion will settle the question of what I am to do in favor of acting as I do. I will call such arguments *practical arguments*, and such conclusions *practical conclusions*. For simplicity, I will say only that my full-blown action corresponds to such a practical conclusion.

My practical reasoning prior to reading the papers was reasoning, in the light of my values, concerning what it would be best to do. I reasoned about what it would be best for me to do as a way of settling the question of what to do. I reached an appropriate practical conclusion which settled this question and acted accordingly. This is the kernel of a plausible account of full-blown action. We may summarize this kernel in two theses:

T1. A full-blown action involves the conclusion of a piece of practical reasoning in which the agent considers reasons he has for so acting. This reasoning, I shall say, *underlies* the full-blown action. The action corresponds to the practical conclusion of the corresponding practical argument. This conclusion settles the question of what to do in favor of this action.[3]

T2. In such practical reasoning one reasons about what it would be best to do, in order to settle the question of what to do. I shall call this *evaluative practical* reasoning.[4]

Now let us consider an apparent case of weak-willed action. Suppose Sam is sitting by a bottle of wine. He knows that if he drinks the wine he will suffer a bad headache in the morning and further that, given the late hour, he must go right to sleep if he is to be fully rested for an important job he must do the next day. Still, he knows that drinking the wine would be quite pleasant and would, temporarily at least, help relieve his present depression.

Sam thinks both that in certain respects his drinking would be, *prima facie*, best, and that in certain other respects his abstaining would be, *prima facie*, best. Weighing these conflicting considerations he concludes that it would be best to abstain, rather than drink. "Still," he thinks to himself while focusing his attention on his reasons for drinking, "the wine does look very good, and it would surely lift me out of my depression." He then reaches for the wine, pours it into a glass and proceeds to drink it.

At this moment his friend drops by and, noting the partially-empty bottle and knowing of Sam's heavy schedule the next day, says to Sam: "Look here. Your reasons for abstaining seem clearly stronger than your reasons for drinking. So how can you have thought that it would be best to drink?" To which Sam replies: "I don't think it would be best to drink. Do you think I'm stupid enough to think that, given how strong my reasons for abstaining are? I think it would be best to abstain. Still, I'm drinking."

Sam drinks the wine for a pair of reasons: the pleasure of drinking and the temporary relief of his depression. He does not drink it impulsively or impetuously; he drinks it, rather, deliberately and with a knowledge of his reasons both for and

against drinking it. Finally, Sam is not compelled so to act. He could put the glass down and go to sleep and he knows this. Sam's case is, it seems, one of free, deliberate, and purposive—i.e., full-blown—action contrary to the agent's best judgment. It seems to be a case of weak-willed action.

On T1 and T2 Sam's action must correspond to an appropriate practical conclusion. What is this practical conclusion? It could not be simply the conclusion that in certain respects his drinking would be, *prima facie*, best; for, on Sam's view, at least that much is also true about his abstaining and so such a conclusion leaves the question of what to do unsettled. The conclusion Sam reaches which does seem to settle the question of what to do is the conclusion that it would be best to abstain. But that cannot be Sam's practical conclusion; for rather than abstain, he drinks. Finally, Sam surely does not also conclude that it would be best to drink; though guilty of some form of irrationality, Sam is not guilty of such blatant inconsistency.

Sam's action must correspond to an appropriate practical conclusion. But the only natural candidate for such a practical conclusion seems to be just his conclusion about what it would be best to do. But this is the conclusion *contrary* to which he must act if his case is to be one of weak-willed action; for it is his acceptance of this conclusion which constitutes his "best judgment" concerning what to do. His purportedly weak-willed action cannot, then, satisfy the demands of T1 and T2 and so cannot really be a case of weak-willed action. The possibility of weak-willed action is ruled out by T1 and T2; or so, at least, it seems.

III

One might try simply to embrace the conclusion of this argument. All apparent cases of weak-willed action would then have to be re-described in one of a variety of ways. In some we could say that the agent was, in fact, compelled to act as he did by some irresistible desire—for example, for drinking the wine. In such cases we cannot directly apply the theory sketched in T1 and T2, so no paradox results. In some other cases we might plausibly insist that though the agent says he thinks option *a* best, his behavior reveals that, at least at the time of action, he thought option *b* best. So no problem arises in applying the framework sketched in T1 and T2.

However, as Sam would surely protest, and as many a cigarette smoker can attest, it seems implausible that every apparent case of weak-willed action will turn out, in this way, to be *merely* an apparent case of weak-willed action. The conclusion of the argument just mooted simply seems wrong.

We must either challenge T1 or T2, or find a way of blocking the inference from T1 and T2 to the conclusion that weak-willed action is not possible. Thesis T1, however, represents an extremely plausible approach to the analysis of full-blown action and its relation to practical thinking. If we are plausibly to reject one of the premises of the argument against the possibility of weak-willed action it will have to be T2.

There is an extreme and a moderate attack on T2. On the extreme attack deliberation about what it would be best to do has no closer relation to practical reasoning than, say, deliberation about what it would be chic to do. If one happens to care about what it would be chic to do, then a consideration of this matter may play an important role in one's practical reasoning. If one does not care, it will be irrelevant. The case is the same with reasoning about what it would be best to do.

On this view, the practical reasoning underlying full-blown action is reasoning about how to do what one desires to do, and these desires themselves neither involve evaluations nor are an essential component of one's evaluations. The argument corresponding to such reasoning might use, for example,

Self, do that which is chic!

as the component in the argument which corresponds to the relevant desire in the reasoning. It would not use any evaluative expression for this purpose. Judgments about what would be best would be relevant only given some further desire to do what would be best.

Thesis T2 would, on this view, be false. Rejecting T2, the defender of this view can say, for example, that practical conclusions are typically of the form

Self, do *a*!

He can then say that there is no more problem with an agent's reaching such a practical conclusion while thinking it best to

do *b* instead than there is a problem with an agent's reaching such a practical conclusion while thinking it chic to do *b* instead.

I shall call this the *extreme externalist* response. This response finds no essential relation between desiring to do something and valuing doing it. It sees judgments about what would be best as playing no special role in the practical reasoning underlying full-blown action.

This response saves the possibility of weak-willed action by severing the connection between evaluation and the practical reasoning underlying full-blown action. This seems to me too high a price to pay. As suggested by our examples, at least one form of practical reasoning underlying full-blown action is reasoning about what it would be best to do. It seems extremely artificial to think of such reasoning as requiring, for its practical force, some further desire in the way in which reasoning about what would be chic requires, for its practical force, such a further desire.

Rejecting the extreme externalist response, one might still offer a moderate criticism of T2. Perhaps some practical reasoning is evaluative practical reasoning; but, still, there will be some cases of full-blown action in which evaluations play no role in the underlying reasonsing. In these cases the underlying reasoning conforms, rather, to the framework sketched by the extreme externalist. If this is correct we can allow for cases of weak-willed action by saying that in *those* cases the underlying reasoning is reasoning of the sort sketched by the extreme externalist. The practical conclusion of such reasoning may then be, for example, of the form

Self, do *a*!

And, again, such a practical conclusion may be reached even though the agent has also concluded that it would be best to do *b* instead.

This more moderate response does not, however, take us very far in understanding how weak-willed action is possible. Consider, again, the case of Sam. His reasoning in favor of drinking conforms to T2. He sees the pleasantness of drinking, and the fact that drinking will relieve his depression, as considerations given which his drinking would be, *prima facie*, the best thing to do. His problem is that he also sees other considerations as making his abstaining, *prima facie*, best; and

he thinks the latter considerations outweigh the former. Still, his reasoning in favor of drinking does seem to be evaluative practical reasoning. So the moderate response—even if it were granted—would still leave it mysterious how Sam's case is possible and so would not get us to the bottom of the problem. To do that we must challenge the cogency of the argument from T1 and T2 to the impossibility of weak-willed action.

IV

The argument under consideration depends on the claim that the only plausible candidate for one's practical conclusion, in a case of evaluative practical reasoning, is one's conclusion concerning what it would be best to do. Since this conclusion also seems to be one's best judgment,[5] there seems to be no room for divergence between best judgment and practical conclusion and so no room for weak-willed action. If we are to block this argument while retaining T1 and T2 we must distinguish an agent's best judgment from his practical conclusion.

One way of doing this would be to allow that the agent's practical conclusion is his conclusion concerning what it would be best to do, but insist that some other relevant evaluative conclusion should be treated as the agent's best judgment. On this approach weak-willed action occurs when these two different conclusions diverge.[6]

But this way of driving a wedge between best judgment and practical conclusion is, I think, flawed. First, if in reasoning about what it would be best to do the agent reaches a conclusion that a certain option would be best, this conclusion will seem the natural candidate for the agent's best judgment. To select any other, distinct conclusion (e.g., that in certain respects a certain option would be, *prima facie*, best) as the agent's best judgment will seem an arbitrary and *ad hoc* way of providing for weak-willed action. Second, this approach will still not be able to provide for Sam's case as originally described. Sam was described as concluding that it would be best to abstain, and yet as still drinking freely, deliberately, and for a reason. Indeed, Sam himself insisted to his friend that this is what he thought. On this approach, however, we cannot let Sam have his way and must insist that, though he does hold *some* evaluative conclusion in favor of abstaining, he thinks that drinking would be best. But this seems wrong.

Finally, there is a further difficulty with the account of full-blown action, used in the argument against the possibility of weak-willed action, which this approach also suffers. That argument tried to show that the practical conclusion corresponding to one's full-blown action, *a*, will be the conclusion that *a* would be best. But one may perform a full-blown action and yet, justifiably, not be at all prepared to draw such a conclusion.

Jones is canoeing downstream and sees that he is approaching a fork in the river. He knows he can neither turn back nor make it to the land, and he knows that the left fork is sunny and the right fork has nice flowers. He also knows that at the end of one and only one of the forks there is a waterfall. Unfortunately, he does not know which fork leads to the waterfall.[7] Jones thinks that there are considerations in favor of each of his options. Suppose, thinking that sunshine is more desirable than flowers, he concludes that with respect to these known considerations it would be, *prima facie*, best to go left and proceeds to act accordingly. Keeping the waterfall in mind, however, Jones might still insist that he is not at all prepared to conclude that going left would be best. Here, it seems, something other than such a conclusion must function as Jones' practical conclusion.

The original argument against weak-willed action was also an argument against describing Jones' case in this way. On this argument we would be obliged, contrary to Jones' insistence, to say of Jones that he has concluded that going left would be best. This seems wrong. The approach we are now considering, however, does nothing to avoid this further difficulty.

To allow for weak-willed action we must, we have seen, drive a wedge between best judgment and practical conclusion. But we cannot plausibly do this by granting that one's practical conclusion is one's conclusion about what would be best, and then finding another candidate for one's best judgment. We need a different account of one's practical conclusion.

V

In evaluative practical reasoning one may reach a conclusion that in certain respects one of one's options would be, *prima facie*, best, and one may also reach a conclusion that one of

one's options would, indeed, be best. The former type of conclusion does not, in general, settle the question of what to do. The latter type of conclusion is not, we have seen, generally present in full-blown action. To solve our problem we need to introduce a further, non-evaluative conclusion to function as the agent's practical conclusion. I propose that we say that practical conclusions are of the form of "I shall do *a*." Such a conclusion would settle the question of what to do. Since it appears that one might conclude

I shall do *a*

while holding that some conflicting option, *b*, would be best, this move promises to allow us to drive the necessary wedge between practical conclusion and best judgment.

But how could one's acceptance of

I shall do *a*

be, strictly speaking, a *conclusion* of reasoning concerning what it would be best to do? Not just anything can serve as a conclusion of such reasoning. There is, it seems, no special problem with treating one's acceptance of

It would be best to do *a*

as a conclusion of such reasoning. Whatever the exact details of the structure of the corresponding practical argument, such a conclusion will involve the same evaluative notion as that involved in the premises. But on the present suggestion this relation will not hold between the premises and practical conclusion of the relevant practical arguments.

I shall call an account which expresses practical conclusions in terms used to express the premises of practical arguments a *homogeneous* account. Both the extreme and the moderate criticism of T2 tried to allow for non-evaluative practical conclusions while retaining a homogeneous theory. This forced them to provide unsatisfactory accounts of the role of evaluations in the practical reasoning underlying full-blown action. The present proposal attempts to allow for non-evaluative practical conclusions while doing justice to the role of evaluations in practical reasoning. Its way of doing this is to adopt a non-homogeneous view of practical reasoning.

In, so to speak, the "psychological mode," the view is that the conclusion, of evaluative practical reasoning, which is involved in full-blown action is not itself a kind of evaluation, but, rather, belongs to a distinct psychological category. What now needs showing is that this is a coherent view.

Here we are faced with an apparent dilemma. On the one hand, we might try to show how

I shall do *a*

can be drawn as a conclusion from

It would be best to do *a*

by providing an account of the latter according to which it entails the former. This is, in effect, Hare's approach.[8] On this approach, however, we will find it difficult not to say that if one really thinks it would be best to do *a* he will accept the cited practical conclusion in favor of *a*. But, we will then be unable to allow for Sam's weak-willed action.[9] Indeed, even if we think it possible for a person to think *a* best, know that this entails a practical conclusion in favor of *a*, and still draw a conflicting practical conclusion instead, it seems implausible that central cases of weak-willed action are ones of such extreme logical error. Further, this approach would still leave it unexplained how Jones' transition to a practical conclusion can be treated as a case of reasoning; for Jones reaches no conclusion concerning what would be best.

On the other hand, if we do not show there to be some relevant norm governing transitions in practical reasoning from the acceptance of certain evaluative conclusions to the acceptance of certain practical conclusions we leave it mysterious how any such transition can be treated as a piece of reasoning.

As a way of steering a path between these two horns, consider the case of decision under risk. In such cases it is a widely accepted principle that the rational thing to do is to choose that option which maximizes one's expected utility. Now, suppose I can choose to pick a ball either from urn A or from urn B. Urn A contains only white balls, which are worth $50. Urn B contains some blue balls, worth $100, and some green balls, worth $1. The ratio of blue balls to green balls in urn B is 1:10. On natural assumptions, I maximize my ex-

pected utility by choosing a ball from urn A. On the cited principle, this is the rational choice to make.

Suppose I decide to take my chances with urn B even though I know that this is to violate the cited principle, in the sense that it is to make what is, on this principle, an irrational choice. A defender of this principle will claim that I am being irrational; but he will not claim that I have failed to deduce a conclusion entailed by premises I accept, or that I am guilty of logical inconsistency.[10] He will say, I take it, that my irrationality does not consist in such extreme logical error but, rather, in violating an acceptable principle of rational choice.

Here we have a principle of rationality in accordance with which one might reason to a practical conclusion without one's premises entailing that practical conclusion. A transition, in evaluative practical reasoning, from the acceptance of certain evaluative conclusions to the acceptance of one's practical conclusion is governed by a principle of this sort.

In evaluative practical reasoning one reaches conclusions of the form of

(a) Given r, a would be, *prima facie*, best[11]

and, sometimes, a further conclusion of the form of

(b) a would be best.

Since both (a) and (b) entail that there is some respect in which a would be good (even if only in its avoidance of something bad), I shall call conclusions of either form *evaluative commitments*.

On T1 and T2, the reasoning underlying a full-blown action, a, will include a consideration of reasons for doing a. So, we may assume that the reasoning underlying a will always issue in the acceptance of at least some evaluative commitment, of the form of (a) or (b), in favor of a. What is needed is some appropriate norm governing the transition from the acceptance of such evaluative commitments to the acceptance of a practical conclusion of the form of

(c) I shall do a.

Sometimes an evaluative commitment will involve considerations which include *both* the considerations involved in

an evaluative commitment in favor of another, conflicting option, *and further* relevant considerations. For example, I might accept

> Given that playing basketball would be fun, it would be, *prima facie*, the best thing to do.

and also accept

> Given that reading the papers would result in missing the fun of playing basketball but would insure that the papers received adequate attention, it would be, *prima facie*, best to read the papers.

Here we could say that the latter evaluative commitment *overrides* the former. If, in addition, I were to reach the conclusion

> Reading the papers would be best.

this, too, would naturally be said to override the former evaluative commitment in favor of playing basketball; for on this last conclusion the balance of *all* relevant considerations weighs in favor of reading the papers.

As a way of generalizing this idea, suppose a person takes mutually conflicting options $a_1, \ldots, a_n$ to be his only options, and considers the evaluative commitments he accepts concerning each of these options. We can define a two-place relation of *overrides* between these evaluative commitments in the following natural way:

(i) "Given r, a_i would be, *prima facie*, best" overrides "Given r', a_j would be, *prima facie*, best" if and only if $i \neq j$, r entails r', and r' does not entail r.

(ii) For any r, "a_i would be best" overrides "Given r, a_j would be, *prima facie*, best" if and only if $i \neq j$.[12]

This notion in hand, it seems plausible to suppose that a rational agent will, on due reflection, accept the following principle:

(PR) It is rational to draw a practical conclusion in favor of a from an accepted evaluative commitment in favor of a unless that evaluative commitment is overridden

> by another evaluative commitment you accept, or would accept if you drew all conclusions entailed by what you already accept.[13]

Transitions from evaluative commitments to practical conclusions are governed by (PR): they are rational or not depending on whether they accord with (PR) or not. Given this norm governing such transitions, the absence of homogeneity does not prevent us from saying that one's evaluative practical reasoning may conclude in one's acceptance of

(c) I shall *a*.

On this view, reasoning from evaluative commitments of the form of

(a) Given *r*, *a* would be, *prima facie*, best

to practical conclusions of the form of (c), will be *defeasible* in character.[14] Even when such an inference accords with (PR) it remains possible for there to be a further premise of the form of

(a′) Given *r* and *r′*, not-*a* would be, *prima facie*, best.

which is consistent with one's original evaluative commitments in favor of *a* but which, if accepted would, on (PR), make the initial inference irrational.

In contrast, an inference from an evaluative commitment of the form of

(b) *a* would be best.

to the appropriate practical conclusion will not be defeasible in this way. There remains no further evaluative commitment which the agent could consistently accept and whose acceptance would, on (PR), make the original inference irrational. Still, one may accept (b) and conclude

(c′) I shall do not-*a*.

without being guilty of logical inconsistency.

On this non-homogeneous account, transitions from evaluative commitments to practical conclusions may be pieces of reasoning. Does this account provide room for, on the one hand, T1 and T2 and, on the other hand, the cases of Jones and Sam?

VI

This account provides a straightforward way of describing Jones, the canoeist. He has drawn—in accordance with (PR)—a practical conclusion in favor of going left from the evaluative commitment

> Given that there is sunshine on the left and flowers on the right, my going left would be, *prima facie*, best.

It is not, on this account, necessary for him to conclude further that his going left would be best.

How can we describe a weak-willed action? A person's full-blown action is contrary to his best judgment if and only if each of his accepted evaluative commitments in favor of what he does is overriden by some other evaluative commitment, which he accepts, in favor of a believed-to-be-conflicting option. A weak-willed action will be contrary to the agent's best judgment in this sense.

Suppose *a* is a weak-willed action. Since it is a full-blown action it will correspond to the practical conclusion

(c) I shall do *a*.

The agent's acceptance of (c) will be the conclusion of evaluative practical reasoning which considered reasons for doing *a*. So, we may assume that in reaching (c) the agent reached some evaluative commitment in favor of *a*. Since *a* is a weak-willed action that evaluative commitment will be of the form of

(a) Given *r*, *a* would be, *prima facie*, best.

In performing *a* the agent reasons from such an evaluative commitment to (c) even though all his accepted evaluative commitments in favor of *a* are overriden by other accepted evaluative commitments in favor of believed-to-be-conflicting options. So, according to (PR), he acts irrationally.[15]

Sam, for example, thinks that in certain respects his drinking would be, *prima facie*, best, but that, when all is taken into account, his abstaining would be best. He has available to him an inference to a practical conclusion which would accord with (PR), namely: an inference from the latter evaluative commitment to a practical conclusion in favor of abstaining. But, it remains possible that he does not, in fact, draw that inference but, rather, focuses his attention on the former evaluative commitment and infers from it a practical conclusion in favor of drinking. Indeed, this is just what Sam does—albeit, irrationally.

On this account weak-willed actions will be rather rare; for when an agent's violations of a principle of rationality like (PR) seem to be widespread and pervasive we can conclude that our system of psychological attributions to that agent is somehow faulty. But it remains possible for one to violate such a principle of rationality. So weak-willed actions will be possible, even given T1 and T2; and that is what we wanted to show.

Against this the following objection might be offered. Granted that the acceptance of a practical conclusion of the form of

I shall do *a*

may be a conclusion of evaluative practical reasoning. Still, in a case of weak-willed action the transition which the agent makes to the acceptance of such a practical conclusion is, *ex hypothesi*, in violation of (PR). Further, we may assume that the agent knows this; for we do not want to make the possibility of weak-willed action depend on the agent's not knowing he is being irrational. But then how could this (faulty) transition be a piece of *reasoning*? Since he knows that it violates (PR) it does not seem that it can be. But if it is not then T1 has not been satisfied.

Let us return to Sam's case. Sam accepts the following evaluative commitments

(i) Given that drinking would be pleasant and relieve my depression, my drinking would be, *prima facie*, best.

(ii) Given that abstaining would allow me to avoid a headache and be rested for tomorrow's job, my abstaining would be, *prima facie*, best.

(iii) Given the considerations cited in both (i) and (ii), my abstaining would be, *prima facie*, best.

(iv) Abstaining would be best.

On the present account Sam's irrationality lies in nonetheless reasoning from (i) to

(v) I shall drink.

The present objection is that this is not a piece of reasoning, but is, rather, a mere transition from accepting (i) to accepting (v).

In reply, note that Sam's transition from (i) to (v) is not a transition from something which, by itself, provides no grounds at all for reaching (v) as one's conclusion. Suppose, for example, Sam knew he had no pressing engagements tomorrow and would not get a headache from drinking the wine. He might, in such a case, justifiably infer (v) from (i). His transition from (i) to (v) might well, in this case, be a case of reasoning.

There are many similarities between our original case and such a case of Sam's justifiably reasoning from (i) to (v). In neither case would it be correct to say that Sam is merely caused to drink by his acceptance of (i); in both cases (i) alludes to the reasons for which Sam drinks. In both cases an explanation of why Sam accepts (v) would have to cite Sam's conception of (i) as providing grounds for (v); in each case his acceptance of (v) depends, in part, on his seeing the transition from (i) to (v) as an instance of a potentially-justified inference pattern. In both cases Sam would, if asked to provide whatever justification he has for drinking, cite the considerations alluded to in (i)—though, of course, in the original case Sam would agree that these considerations do not suffice fully to justify his action.

I claim that these features, which Sam's original case shares with the case of justified reasoning from (i) to (v), suffice to make the original case one of reasoning. Sam's irrationality in the original case does not preclude it from being a case of reasoning but is, rather, a feature of that

reasoning. We might try to explain why Sam reasons in this admittedly irrational way by, for example, appealing to the depth of his depression, the strength of his desire to get out of it, and, perhaps, to the way in which his depression dulled his appreciation of his reasons for abstaining. But however we explain Sam's case, *what* we are explaining is his drawing (v) as a conclusion from (i).

VII

On T1 and T2 full-blown action involves the conclusion of a piece of evaluative practical reasoning. This conception rejects extreme externalism: on T1 and T2 one's evaluations play a crucial role in the reasoning underlying full-blown action. In contrast, T1 and T2 together do not oblige us to accept what might naturally be called "extreme internalism": the view that, in addition, full-blown action corresponds to the evaluative conclusion of the relevant practical argument which settles the question of what to do, or, at least, to a conclusion entailed by that evaluative conclusion. Our non-homogeneous account allows for the essential role of evaluations in practical reasoning while insisting on the need for a further step—governed by (PR)—from the evaluative conclusions of such reasoning to the conclusion involved in acting. This is the position in favor of which the phenomenon of weak-willed action seems to argue.

References

[1] Aristotle, *De Motu Animalium*.
[2] Bruce Aune, *Reason and Action* (Dordrecht-Holland: Reidel, 1977).
[3] Robert Binkley, "The Logic of Action," in Myles Brand and Douglas Walton, eds., *Action Theory* (Dordrecht-Holland: Reidel, 1976): 87-104.
[4] Hector-Neri Castañeda, *Thinking and Doing* (Dordrecht-Holland: Reidel, 1975).
[5] Roderick Chisholm, "The Ethics of Requirement," *American Philosophical Quarterly* 1(1969): 147-53.
[6] Donald Davidson, "How is Weakness of the Will Possible?" in Joel Feinberg, ed., *Moral Concepts* (Oxford: Oxford Univ. Press, 1969): 93-113.
[7] P. T. Geach, "Dr. Kenny on Practical Inference," *Analysis* 26(1965-66): 76-9.
[8] R. M. Hare, *The Language of Morals* (Oxford: Oxford Univ. Press, 1952).
[9] Plato, *Georgias*
[10] Wilfrid Sellars, "Thought and Action," In Keith Lehrer, ed., *Freedom and Determinism* (New York: Random House, 1966): 105-39.

NOTES

[1]An earlier version of this paper was presented at a colloquim at the California State University at Hayward, and at a symposium on Action and Responsibility at the California State College at Sonoma. In writing this paper I was influenced by Donald Davidson's fascinating discussion of similar problems in [6]. I am also very much indebted both to Donald Davidson and to John Perry for many rewarding conversations on the topic of this paper. Others to whom I am indebted include Robert Audi, David Charles, Keith Donnellan and the anonymous reviewers for *Noûs*.

[2]In [4] Castañeda holds that this structure will essentially include "practitions" which are similar to, but distinct from propositions. In this paper I side-step the question of whether he is right about this matter.

[3]A version of T1 is suggested by Aristotle in [1]: 701a14-21. Other philosophers who seem to hold a version of T1 include Robert Binkley in [3], Hector-Neri Castañeda in [4], Donald Davidson in [6], and Wilfrid Sellars in [10].

[4]This Socratic view (see [9]: 467-8) is suggested by Davidson in [6].

[5]Strictly speaking, the agent's best judgment is his state of valuing certain of his options in certain ways. However, here and elsewhere I take the liberty of speaking of the expression or content of that state as his best judgment, and depend on the context to disambiguate.

[6]This is, in effect, the proposal made by Davidson in [6]. On Davidson's view, a weak-willed agent's best judgment is his view of which option would be best, *all things considered*. In contrast, his practical conclusion says which option would be best. On Davidson's analysis these are two distinct conclusions. Against Davidson's proposal I would offer versions of each of the three arguments in the text against this approach.

[7]Here I was helped by David Pears.

[8]See [8]: Chapter 11. Hare would use "Do a" as the relevant practical conclusion.

[9]Castañeda, in [4], seems to have analogous difficulties. For Castañeda, the practical conclusion in favor of a is a first-person practition to do a. This practical conclusion is distinct from a conclusion that one "ought overridingly to do a"—where this latter conclusion, I take it, is the agent's best judgment. So, on Castañeda's account, as on mine, one's best judgment is distinct from one's paratical conclusion. But Castañeda goes on to understand "I ought overridingly to do a" in such a way that "to think firmly that one ought overridingly to do a is in fact to endorse the practition to do a" (p. 304—I have changed the variables). But this appears to rule out the possibility of cases of weak-willed action along the lines of Sam's case.

[10]Bruce Aune makes a similar point (but develops it differently) in [2]: 137-42.

[11]I understand this to entail r.

[12]Davidson discusses a similar notion in [6]: 111 as does Chisholm in [5]: 148. In formulating (i) I assume that r and r' include only considerations taken by the agent to be relevant.

[13]I assume that the agent does not hold logically inconsistent views. Note that one might, according to (PR), be rational, given one's grounds, in drawing the practical conclusion one draws, but still be faulted for not looking carefully enough for considerations against this conclusion.

[14]For helpful discussions of defeasibility as a feature of practical reasoning see [5] and [7].

[15]I assume that if it is not rational, on (PR), to reach a certain conclusion from certain grounds it is irrational to reach that conclusion from those grounds.