

Weakness of the will and the appearance of the good

PHIL 43503
2/9/08

Intentional action and appearances of the good	1
Internalism and externalism	2
Weakness of the will and strong appearance of the good theories	3
<i>Arguments for the impossibility of weakness of the will</i>	3
<i>Plato on weakness of the will</i>	3
<i>The argument from judgement internalism</i>	4
<i>Examples of weakness of the will</i>	5
<i>Defending the strong appearance of the good theory</i>	6
<i>Making room for (semi-)akrasia</i>	7
<i>Distinguishing between akrasia and compulsion</i>	7
<i>Theoretical akrasia</i>	8
<i>The impossibility of global akrasia</i>	8
Weak appearance of the good theories and extreme weakness of the will	8

Intentional action and appearances of the good

We are interested in a cluster of theories of intentional action which can be seen as alternatives to the various other causal theories we have so far discussed. These are theories that say that what it is to Φ intentionally is for one's Φ ing to be caused by Φ appearing as a good, or the best, course of action. In particular, we distinguished between the following two versions of this sort of theory:

The strong appearance of the good theory

A Φ s intentionally =df A's Φ ing is caused by the judgement that A has most reason to Φ / Φ ing is best thing for A to do

However, examples of weakness of the will are not counterexamples to the following sort of appearance of the good theory, which we discussed in connection with Aquinas:

The weak appearance of the good theory

A Φ s intentionally =df A's Φ ing is caused by A's judgement that Φ ing is good/ Φ appearing as good to A

For these theories to be true, two things must be the case:

- (1) It must be the case that whenever one's Φ ing is caused by one seeing Φ as a good, or the best, course of action, one's Φ ing is intentional.
- (2) It must be the case that whenever one intentionally Φ s, one sees Φ ing as a good, or the best, course of action.

These are the two "directions" of the above biconditionals.

Internalism and externalism

To be sure that (1) is true, we need to rule out the following sort of scenario: A judges that Φ ing is best, and this caused him to Φ , but he didn't do so intentionally; on the contrary, he intended to do something quite different.

This would be in a way the opposite of weakness of the will: rather than intentionally acting contrary to your best judgement, you would be acting according to your best judgement, but not intentionally.

One way to rule out this sort of possibility would be to endorse an internalist thesis of the following sort:

Necessarily, if A judges that she has most reason to Φ (that Φ ing is the best thing to do, ...), then A intends to Φ

Let's use "externalism" as a name for the denial of this thesis, i.e.

Possibly, someone judges that she has most reason to Φ , and does not intend to Φ

Both these labels can be and are used to stand for many different theses, but this is what they'll mean for our purposes today.

There is an important connection between these theses and the above theories of intentional action. Suppose that externalism is true. Then, since someone can believe that he has most reason to Φ without intending to Φ , it is plausible that he can believe that he has most reason to Φ while intending to do something quite different. But then if he succeeds in doing this something quite different, that action will not be accompanied by any judgement about the good of the action. That would be a problem for at least some appearance of the good theories of intentional action.

This shows that if externalism is true, then it seems likely that at least strong appearance of the good theories of intentional action are false (i.e., these theories entail the truth of internalism).

Weakness of the will and strong appearance of the good theories

Claim (2) above -- that intentional action is always accompanied by a judgement that the action is a good, or the best, course of action -- is threatened by the possibility of weakness of the will.

Let's say that an act is an instance of weakness of the will if and only if (1) the agent did it intentionally and (2) the agent judged at the time of the act that it was not the best course of action, all things considered, for him to do. If there are cases of this sort, then the strong appearance of the good theory is false.

Arguments for the impossibility of weakness of the will

Plato on weakness of the will

There is a long tradition of denying that weakness of the will, so described, is possible. A locus classicus for this view is Plato's *Protagoras*, in which Socrates seems to argue that weakness of the will is impossible.

Here is the central passage in that argument:

Having noted this result, suppose we reinstate the names pleasant and painful for the same phenomena, thus: A man does—*evil* we

said before, but now we shall say *painful* actions, knowing them to be painful, because overcome by pleasures—pleasures, obviously, which were not a match for the pains. And what meaning can we attach to the phrase *not a match for*, when used of pleasure in relation to pain, except the excess or deficiency of one as compared with the other? It depends on whether one is greater or smaller, more or less intense than the other. If anyone objects that there is a great difference between present pleasure and pleasure or pain in the future, I shall reply that the difference cannot be one of anything else but pleasure and pain. So like an expert in weighing, put the pleasures and the pains together, set both the near and distant in the balance, and say which is the greater quantity. In weighing pleasures against pleasures, one must always choose the greater and the more; in weighing pains against pains, the smaller and the less; whereas in weighing pleasures against pains, if the pleasures exceed the pains, whether the distant, the near, or vice versa, one must take the course which brings those pleasures; but if the pains outweigh the pleasures, avoid it. Is this not so, good people?' I should say, and I am sure they could not deny it.

Protagoras agreed.

'That being so then, answer me this, I shall go on. The same magnitudes seem greater to the eye from near at hand than they do from a distance. This is true of thickness and also of number; and sounds of equal loudness seem greater near at hand than at a distance. If now our happiness consisted in doing, I mean in choosing, greater lengths and avoiding smaller, where would lie salvation? In the art of measurement or in the impression made by appearances? Haven't we seen that the appearance leads us astray and throws us into confusion so that in our actions and our choices between great and small we are constantly accepting and rejecting the same things, whereas the metric art would have canceled the effect of the impression, and by revealing the true state of affairs would have caused the soul to live in peace and quiet and abide in the truth, thus saving our life?' Faced with these considerations, would people agree that our salvation would lie in the art of measurement?

How is this argument supposed to work? Does the argument assume the truth of psychological egoism?

Socrates states the conclusion of his argument as follows:

Then if the pleasant is the good, no one who either knows or believes that there is another possible course of action, better than the one he is following, will ever continue on his present course when he might choose the better. To 'act beneath yourself' is the result of pure ignorance; to 'be your own master' is wisdom.

The argument from judgement internalism

Let *judgement internalism* be the view that there is an internal -- i.e., necessary -- connection between the judgement that some action is good (or the thing that I ought to do) and the motivation to perform that act. If this view is true, it is impossible to judge that an act is good without being motivated to perform that act.

This idea that there is a necessary connection between evaluative judgement and motivation might suggest that there is also a necessary connection between comparative evaluative judgements and the strength of one's motivations:

if one judges that A is better all things considered than B, one has a stronger motivation to do A than B.

But the following claim also looks plausible:

when one acts intentionally, one always does what one has the strongest motivation to do.

Do these principles rule out the possibility of weakness of the will?

Can you accept judgement internalism while denying either of these principles?

Examples of weakness of the will

Despite the arguments for the impossibility of weakness of the will, plausible examples of the phenomenon are not hard to come by. Stocker gives examples of intentional actions of two types:

- (1) Cases in which an agent fails to be motivated to perform an action, despite believing it to be good.
- (2) Cases in which an agent performs an action because of bad they see in the action.

Suppose that these cases are genuine. Do they show that weakness of the will is possible?

Stocker's first example of (1) is the example of the politician. Filling out this case, he says

Suppose it is because of bitterness at the way the politician was treated that he does not desire to help those people. He has ceased caring about or for them. Perhaps he dislikes them. His non-attraction—his indifference or hostility—to the (believed) good confutes the thesis that the (believed) good must attract.

One might reply to this case by trying to find some perceived good in the politician's inaction, such as the goodness of the peace of mind brought on by his inaction. In reply, Stocker says:

This objection to my claim is problematic, however, whether the thesis is taken in a comparative or noncomparative form. Rejecting its first suggestion, I would argue for the following: what the politician wants can be simply that those people not be helped by him or that they not have that good. Dislike or bitterness or not caring for or about are all sufficient explanations of such non-attraction to the good of someone. They need not be supplemented by some other state or condition, in particular some egoistic state or condition, to make the non-attraction intelligible. To be sure, each of these replies needs further discussion. But for reasons concerning the second suggestion, we need not pursue them.⁹

Another example of type (1): depression.

A case of type (2): desire for junk food on the basis of its being bad for me; desire to harm others.

Another well-known example of type (2) can be found in Book 2 of Augustine's *Confessions*:

“Yet I lusted to thief, and did it, compelled by no hunger, nor poverty, but through a cloyedness of well-doing, and a pamperedness of iniquity. For I stole that, of which I had enough, and much better. Nor cared I to enjoy what I stole, but joyed in the theft and sin itself. A pear tree there was near our vineyard, laden with fruit, tempting neither for colour nor taste. To shake and rob this, some lewd young fellows of us went, late one night (having according to our pestilent custom prolonged our sports in the streets till then), and took huge loads, not for our eating, but to fling to the very hogs, having only tasted them. And this, but to do what we liked only, because it was disliked. Behold my heart, O God, behold my heart, which Thou hadst pity upon in the bottom of the bottomless pit. Now, behold, let my heart tell Thee what it sought there, that I should be gratuitously evil, having no temptation to ill, but the ill itself. It was foul, and I loved it; I loved to perish, I loved mine own fault, not that for which I was faulty, but my fault itself.”

Suppose that Stocker is right that there are genuine examples of cases of types (1) and (2). How would you argue on that basis that weakness of the will, as defined above, is a genuine phenomenon?

A natural response to Stocker's argument is externalism: the view that there is no necessary connection at all between evaluative judgements and intentions or motivation (though there may, of course, be systematic connections between the two in the case of particular agents). As noted above, this looks like a problem for at least the strong appearance of the good theory.

Defending the strong appearance of the good theory

Appearance of the good theorists have not simply given up in response to these sorts of examples; they've responded by both trying to explain the sorts of cases Stocker emphasizes, and by arguing that externalists give an incorrect account of akrasia.

Making room for (semi-)akrasia

Typically, internalists will deny the existence of akrasia as characterized above, and try to explain seeming examples of this sort of weakness of the will as genuine examples of something else. One strategy for doing so, which one finds in Davidson as well as Tenenbaum, is to distinguish the following two phenomena:

Unconditional or all-out judgement

Judging that Φ is what I have most reason to do./Judging that Φ is best.

All things considered judgement

Judging that Φ is what I have, all things considered, most reason to do./Judging that Φ is, all things considered, best.

Weakness of the will, as defined above, is action contrary to unconditional judgement, and internalists must deny that this is possible. But they say that seeming examples of this sort of weakness of the will are really just examples of acting contrary to one's all things considered judgement -- which is consistent with externalism, and appearance of the good theories.

Is this an adequate response to Stocker's examples?

An objection: would the akratic agent agree if you told him that he did not unconditionally judge that he was doing what he did not have most reason to do? Can the internalist simply say that the akratic agent is wrong about his own judgements?

Distinguishing between akrasia and compulsion

This sort of view of weakness of the will also brings out an important difference between how internalists and externalists think of weakness of the will. Internalists who take the route described above think of examples of akrasia as cases in which there is a certain sort of cognitive conflict within the subject: a conflict between two sorts of judgements.

Externalists, on the other hand, think of cases of akrasia as cases in which there is a conflict between judgement, or evaluation, and *motivation*. This is the natural view for externalists, who think of judgement and motivation as only contingently connected.

But this poses a problem for externalists: if cases of akrasia are genuinely cases in which one's desires or other motivating states overcomes one's judgements, what is the difference between akrasia and compulsion? What in short, makes our examples of akrasia genuine intentional actions, as they appear to be?

This is a question to which the internalist has a ready answer: cases of akrasia, but not cases of compulsion, are caused by unconditional evaluative judgements.

Possible reply: to say that the externalist treatment of weakness of the will does not explain the distinction between akrasia and compulsion does not imply that the externalist cannot explain this distinction in some other way.

Theoretical akrasia

This is connected to an interesting defense of internalism from Tenenbaum. He suggests that there are examples of akrasia in the realm of theoretical reason -- belief formation -- as well as practical reason -- reasoning about what to do.

The example of skepticism about the external world.

He thinks that in these cases, while we think that we have, all things considered, most reason to believe p , a certain other claim, q , which is inconsistent with p , just strikes us as true (appears to us as true). Sometimes in such cases we form the belief in q , against our all-things-considered best judgement. This is parallel to the internalist's treatment of akrasia in the practical case, which seems good if they are instances of the same phenomenon.

Moreover, Tenenbaum claims that there is no way to explain these cases in terms of evaluation and motivation coming apart -- is he right about this?

Possible reply: say that a conflict between motivation and best judgement is also what is going on in the theoretical case.

The impossibility of global akrasia

In a paper which we did not read, Sarah Stroud argues that by divorcing judgement and intention, externalists must accept the possibility of global akrasia: a world in which there are many intentional agents, but no one ever acts according to their judgement about what they have most reason to do. Two questions: (i) Is externalism committed to the possibility of such a world? (ii) Is such a world really impossible?

Weak appearance of the good theories and extreme weakness of the will

Cases of weakness of the will do not as such raise a problem for weak appearance of the good theories.

Aquinas is a possible example of someone who holds such a view. According to Aquinas, the nature of the will is such that it can only be moved by the apparent good -- i.e., by things which the agent "apprehends", or takes to be, good:

But it must be noted that, since every inclination results from a form, the natural appetite results from a form existing in the nature of things: while the sensitive appetite, as also the intellective or rational appetite, which we call the will, follows from an apprehended form. Therefore, just as the natural appetite tends to good existing in a thing; so the animal or voluntary appetite tends to a good which is apprehended. Consequently, in order that the will tend to anything, it is requisite, not that this be good in very truth, but that it be apprehended as good. Wherefore the Philosopher says (*Phys. ii. 3*) that *the end is a good, or an apparent good*.

Note that this conclusion is a bit weaker than Plato's: the idea is not that no one ever does anything which he does not take to be the best course of action all things considered, but rather that no one ever does anything which he does not take to be good to some extent.

So it seems that Aquinas needn't rule out the possibility of weakness of the will. Rather, it seems that he is committed only to ruling out what we might call extreme weakness of the will: intentionally performing an act despite judging at the time of the act that there is *nothing* good about it at all.

Are any of Stocker's cases plausible examples of weakness of the will?

Are there any disadvantages of the weak appearance of the good theory over the strong?