

# WILEY

## International Phenomenological Society Philosophy and Phenomenological Research

---

Dispositions and Fetishes: Externalist Models of Moral Motivation

Author(s): James Dreier

Source: *Philosophy and Phenomenological Research*, Vol. 61, No. 3 (Nov., 2000), pp. 619-638

Published by: [International Phenomenological Society](#)

Stable URL: <http://www.jstor.org/stable/2653615>

Accessed: 05-11-2015 16:39 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Wiley, International Phenomenological Society and *Philosophy and Phenomenological Research* are collaborating with JSTOR to digitize, preserve and extend access to *Philosophy and Phenomenological Research*.

<http://www.jstor.org>

# Dispositions and Fetishes: Externalist Models of Moral Motivation\*

JAMES DREIER  
*Brown University*

Internalism says that if an agent judges that it is right for her to  $\phi$ , then she is motivated to  $\phi$ . The disagreement between Internalists and Externalists runs deep, and it lingers even in the face of clever intuition pumps. An argument in Michael Smith's *The Moral Problem* seeks some leverage against Externalism from a point within normative theory. Smith argues by dilemma: Externalists either fail to explain why motivation tracks moral judgment in a good moral agent or they attribute a kind of fetishism to good moral agents. I argue that there are alternative models of moral motivation available to Externalists, in particular a model according to which a good moral agent is one who is effectively regulated by a second order desire to desire to do what is right.

## 1. Introduction

Here is a well known thesis about the relation between moral judgment and motivation:

(Internalism) If an agent judges that it is right for her to  $\phi$ , then she is motivated to  $\phi$ .<sup>1</sup>

Internalism is presented as a conceptual truth about moral judgment, not as a substantive moral thesis. It is at issue in *metaethics*, rather than in normative theory proper. Not a metaethical theory itself, Internalism is an alleged fact about ordinary moral thought, a datum for which a metaethical theory must account. If it is a fact, then a theory gains an advantage insofar as it explains Internalism, and loses ground if it is inconsistent with Internalism. But metaethicists do not agree about Internalism. Those who take it to be a fact about ordinary moral thinking are Internalists. Others are Externalists. Exter-

---

\* Thanks to David Schmitz for some comments on a draft; to an anonymous referee for two useful suggestions; and to Sarah Wright for some insightful remarks about the general subject that she made in a seminar. Sarah's comments got me thinking along the lines presented in this paper.

<sup>1</sup> This is one version of Internalism, a simplified derivative of the one Michael Smith uses in (Smith 1994); see pp. 60–61, and for other versions see the works cited on those pages.

nalists, it appears, simply don't share the intuitions that seem so powerful to Internalists.<sup>2</sup>

The disagreement between Internalists and Externalists runs deep, and it lingers even in the face of clever intuition pumps.<sup>3</sup> This debate in metaethics might be at a standoff, each side standing fast on its intuitions. Standoffs of this sort in philosophy are depressing.

In "Persons, Character, and Morality" Bernard Williams told an influential story whose point was intended to be a point inside of normative theory proper.<sup>4</sup> Williams considers a discussion of Charles Fried's<sup>5</sup> in which a man could "save one of two persons in equal peril, and one of those in peril was, say, his wife." Fried remarks that "surely it would be absurd to insist that ... he must treat both equally, perhaps by flipping a coin." Williams adds, "surely *this* is a justification on behalf of the rescuer, that the person he chose to rescue was his wife?" And he says that by this something "ambitious" might be intended,

essentially involving the idea that moral principle can legitimate his preference, yielding the conclusion that in situations of this kind it is at least all right (morally permissible) to save one's wife.... But this construction provides the agent with one thought too many: it might have been hoped by some (for instance, by his wife) that his motivating thought, fully spelled out, would be the thought that it was his wife, not that it was his wife and that in situations of this kind it is permissible to save one's wife.<sup>6</sup>

I will take it that the lesson to draw from the example is that there is something objectionable about a person who is motivated by the thought that saving his wife is the right (or permissible) thing to do, rather than by the thought that he can only save one and that woman is his wife. What "might

---

<sup>2</sup> For example, in (Brink 1986), (Brink 1989), esp. chapter three; and more recently, (Brink 1997).

<sup>3</sup> As (Dennett 1980) calls them. The pumps I'm thinking of are Hare's Missionaries, in (Hare 1952), 9.4, pp. 148–50, and also Horgan's and Timmons' Moral Twin Earth, in (Horgan 1992). Hare imagines a missionary captured by cannibals who intend to eat him. The missionary attempts to persuade the cannibals that it would be wrong to eat him. At first they do not understand what he means by 'wrong'; the missionary explains by telling them what sorts of things are called 'wrong'. The cannibals come to understand the descriptive content, but wonder why the wrongness of eating a missionary should count *against* it.

Horgan's and Timmons' Moral Twin Earth is populated by English speakers who subscribe to a broadly utilitarian moral system in reflective equilibrium, so that although they agree with earthling Anglophones in many moral judgments, they disagree in some significant ways. Horgan and Timmons claim that the disagreement does not show that the twin-English word 'wrong' merely has a somewhat different meaning from the English word; we are still inclined to think that we and the Twin Earthlings *disagree* about something, that our dispute isn't merely semantic.

<sup>4</sup> (Williams 1981).

<sup>5</sup> (Fried 1970), p. 227.

<sup>6</sup> (Williams 1981), p. 18.

have been hoped” was that the man would be motivated by a desire to save his wife, and not by a desire to do the right thing.

Intuitions about Williams’ example are strong, and they appear to be independent of metaethical theory. In *The Moral Problem*, though, Michael Smith pressed Williams’ point into service on behalf of Internalism, against Externalism, in what I will call (following Hallvard Lillehammer<sup>7</sup>) the Fetishism argument. Smith presents this argument as a kind of dilemma, but Lillehammer focuses on one horn of the dilemma, the “fetishism horn”. Here is Smith’s argument.<sup>8</sup>

A good and strong-willed person will be reliably motivated to do what he believes to be right. Each metaethical theory must explain this fact. Internalist metaethical theories will already have an explanation, for any Internalist theory has already explained a conceptual connection between moral judgment and motivation. Externalists have no explanation of any such conceptual connection, because they do not believe that there is one. But they too must clearly agree that there is a conceptual connection between being a *good*<sup>9</sup> *moral agent* and being reliably motivated to do what is right. Externalists will explain this connection by saying what it is to be a good moral agent. Smith thinks that they have to say that a good moral agent is one with good moral motivations, and that in one or another way this characterization will be cashed out by saying that a good moral agent is one who is motivated to do what is right. Now consider the following sentence attributing moral motivation to Kalista:

(K) Kalista desires to do what is right.

(K) is ambiguous. It could mean that for each thing that is in fact right, Kalista desires to do that thing. Or it could mean that Kalista has a desire whose content is: to do whatever is right. Smith calls the first reading the *de re* reading, and the second the *de dicto* reading. The argument against Externalism is a dilemma. If the Externalist chooses the *de dicto* reading he is impaled on one horn, and if he chooses the *de re* reading then he is impaled on the other.

Suppose (K) is given the *de re* reading. Kalista is a good moral agent in virtue of her desiring to do what is, as a matter of fact, right. Now clearly she

---

<sup>7</sup> (Lillehammer 1997).

<sup>8</sup> (Smith 1994), pp. 71–76. In what follows I ignore some technical difficulties in sorting out distinctions between desires to do what is right and desires to do what one believes to be right. Some of these are discussed by Lillehammer. I don’t think it is important to the argument to spell out the different sorts explicitly, and to do so would encumber the exposition to follow.

<sup>9</sup> Good and strong-willed, that is. Hereafter I omit the ‘strong-willed’. In (Smith 1997), Smith uses ‘moralist’ instead of ‘good and strong-willed person’, and at p. 111, n. 27, he explains that he regrets using the earlier terminology which he now thinks is misleading.

won't be a good moral agent if her motivation is, say, to get a promotion by pleasing her fine upstanding boss. Rather she must be motivated *non-derivatively* as Smith says, or as I shall say, *originally*, by whatever features of actions make them right. But now suppose that Kalista is at first a libertarian, but comes to be convinced that libertarianism is incorrect and that a more utilitarian moral theory is correct. Then good moral agent that she is, she will stop voting for Libertarian political candidates and start voting for, say, Social Democrats. A good moral agent's motivations track her moral beliefs. Call this the *tracking condition*. The *de re* interpretation of (K) cannot explain the tracking condition. For original concern for certain values is not apt to change in the wake of changing belief. The *de re* reading is impaled on the tracking horn.

Suppose (K) is instead given the *de dicto* reading. Kalista is a good moral agent in virtue of her desiring to do whatever turns out to be the right thing to do. This desire is one she could have even if she has no idea of what the right thing to do is, or if she is uncertain. On the *de dicto* reading, (K) does explain tracking. But if the *de dicto* reading of (K) does characterize Kalista, then she appears to be the sort of agent that Williams found unattractive. She is motivated to vote for Social Democrats, or to give money to famine relief, not by the thought that in so doing she will benefit the starving and the needy, but by the thought that in so doing she will do what she is morally required to do. And this motivation is *not* what characterizes a good moral agent. It is *fetishistic*, focusing on an aspect of the action removed from the aspect that motivates good moral agents.<sup>10</sup> So the *de dicto* reading is impaled on the fetishism horn.

If Smith's argument is successful, then the standoff in metaethics might be broken. It would, at least, draw support for Internalism from some more squarely normative intuitions, and that would be very welcome help. Unlike Lillehammer, I do think that the Fetishism argument is at least largely successful against the *desire de dicto* Externalist explanation of moral motivation.<sup>11</sup> But this is far from the end of the story. For there are Externalist alternatives.

---

<sup>10</sup> What Smith calls "fetishism" may be what J. J. C. Smart has in mind when he says that deontological theories are infected with "rule worship". See (Smart 1973), p. 6.

<sup>11</sup> I say "largely" because while I think that a concern with rightness *per se* does look morally unattractive when the object of moral concern is a particular person and her well-being, and especially so when the person bears some special morally significant relationship to the deliberator, the unattractiveness seems to me to dissipate somewhat when the object of moral concern is something more abstract like equality, or fairness, or fidelity to promises. Is it so ugly for a promisor to use the rightness of promise-keeping as his reason for honoring a promise, rather than having an original concern for keeping promises? I suppose I do think it is more admirable to be motivated by a direct concern for promise-keeping, or perhaps for whatever it is that is important about promise-keeping, rather than for promise-keeping's rightness. But it's harder to get excited about it

In section 2, I present one of the alternatives: what I call the model of Suggestible People. I consider two objections to this model and answer them. A fuller answer to the second objection occupies section 3, where I explain two different ways in which a desire can be conditional. In section 4, I present a different alternative Externalist model: the model of a person with the second order desire to desire to do whatever is in fact right. I also argue (a) that this model is distinct from the model of Suggestible People, and (b) that it is superior to the model of Suggestible People. In section 5, I reply to an objection to the second order desire model, namely, that it suffers from its own sort of fetishism. I conclude that there is no sound argument from Williams' point to Internalism.

## 2. A Third Externalist Model: Suggestible People

An analogy will help to clarify Smith's argument, and also to suggest the third path for Externalists.

### 2.1 My List of Foods

Suppose that I kept a list of foods, changing my list from time to time and giving you hints about what was on the list, occasionally showing it to you. And suppose that you had a desire to eat whatever was listed. There are two ways you could want to eat what was on my list. First, it might just so happen that I have listed exactly those foods that you like to eat. Or maybe it isn't a coincidence, maybe I have been keeping tabs on you, and my list is called "Things you like to eat". Then your desire would be *de re*: it is true of each item on my list that you desire to eat it. Admittedly this would be an odd thing for me to do. But second, you might be the odd one. You might have a peculiar desire *de dicto*, whose content was: to eat whatever is on the list. Then you would ask me what was listed today and try to sneak a peak every now and then. If you saw me writing 'avocado', you would want to eat some avocados, even though at the moment you can't stand them.

Now clearly in the first case you would not meet the tracking condition. As my list changed, your culinary desires would not change in step with it (unless by amazing coincidence, or if I were trying to list your favorite foods and kept good track). The *de re* desire (or really, desires, since it would be at best very unnatural to say that you had a single desire: to eat those things which are, as a matter of fact, on my list; we would surely say that you had a bunch of desires, one desire to eat avocados, one desire to eat banana chips, one to eat cucumbers, and so on) does not meet the tracking condition.

The problem with the *de re* desire to do what is right is precisely that it does not meet the tracking condition. A good moral agent will meet the track-

---

than it is to scorn the man who is motivated by the thought that it is morally right to save one's wife. Felicia Ackerman pointed this out to me in conversation.

ing condition; her motivations will change to keep in step with her moral beliefs.

In the second case, you would clearly meet the tracking condition. Insofar as your beliefs about my list were correct, your motivations would change with my list, you would continue over time to desire (*de dicto*) to eat what was listed. And even when your beliefs were false, you would still desire to eat what you believed to be on my list, and your motivations would track your beliefs. In this respect, you would be like the second Externalist model of the good moral agent, whose motivations track what she believes to be good, even as those beliefs change.

But when you came to desire to eat avocados, a true avocadophile would frown, for you would not want to eat avocados for the right reason. You would want to eat them only because they are listed. The avocadophile thinks this is a disgusting attitude, a kind of list fetish. And the second Externalist model of the good and strong-willed moral agent is subject to a fetishism objection, too. A good moral agent wants to relieve the suffering of others, not because relieving suffering is right, not because it is in the extension of 'right', not because it is on the List of Right Actions, but because she cares directly for those who suffer. She likes avocados.

Can we imagine a person who meets the tracking condition and also avoids fetishism?

## 2.2 *Suggestibility Explained*

Suppose you have no particular desire (*de dicto*) to eat what is listed, but you do have an equally odd disposition: you are disposed to desire all and only those foods you come to believe are on my list. Maybe you wear a patch on your arm that delivers a slow but steady drip of medication that maintains this disposition by exploiting your endocrine system. To you it feels like this: you can't stand oysters, but as you watch me inscribe 'oysters' on my list, you suddenly warm to the idea, and you develop a craving for them. Call a person with such a disposition, *suggestible*.

A suggestible person wants to eat whatever she believes is on my list. She does not have any desire *de dicto* with the content: to eat what is on my list. And when she comes to desire avocados, she does so in a way that an avocadophile admires, for she just loves them. The suggestible person wants avocados (when she sees them named on my list) in as thoroughly non-instrumental a way as can be. So she is no fetishist. But obviously she does meet the tracking condition.

The *morally suggestible* person has a (structurally) similar disposition. It is a disposition to want to perform a certain kind of action upon coming to believe that that kind of action is morally right. The morally suggestible person has no standing *de dicto* desire to do what is right, or what she

believes is right, and when she does desire to do certain actions which she believes right, she doesn't desire them in a fetishistic way. She really does love avocados. And she meets the tracking condition. Because of her suggestibility, she predictably and regularly wants to do whatever she believes to be right.

David Copp has proposed something very like this model as an alternative to the two Externalist models suggested and rejected by Smith. Copp writes,

[C]onsider the idea of a disposition to desire straightaway to do what one believes to be right. A good and strong-willed person might have this disposition. If so, and if she comes to believe it is right to vote, she desires straightaway to vote without deriving this desire from an underlying desire.<sup>12</sup>

It is this sort of person that I have in mind. I am calling such a person, a morally suggestible person. A morally suggestible person meets both challenges: she meets the tracking condition, and she is no fetishist. Can an Externalist use moral suggestibility as a model of good moral agency? Or is there something wrong with the model?

### 2.3 *Two Objections*

#### First Objection

The description of suggestible people is a sham, because it attributes to such people certain dispositions while denying that they have certain *de dicto* desires. The dispositions attributed simply *are* the *de dicto* desires in question. So it's as though you have attempted to stipulate that a certain polygon has three sides but not three angles.

I think this objection may be one that Michael Smith used in *The Moral Problem*, though not as an objection so much as a hasty way to conclude that on an Externalist model, a good moral person *must* be one who desires to do what is right. He attributes to the Externalist the claim "that what explains the reliable connection between judgement and motivation is a motivational disposition I have in virtue of which I count as a good person."<sup>13</sup> Smith adds, "In other words, what explains the reliability of the connection is the *content of my moral motivation*." And then he canvasses the two possible motivational contents: the content given by the *de re* reading of (K) and the one given by the *de dicto* reading of (K). Later, Smith puts what I take to be the same point like this.

[Externalists could say] that moralists change their motivations in this way because what makes someone a moralist is the fact that they are simply so disposed that they change their desires

---

<sup>12</sup> (Copp 1997), p. 50.

<sup>13</sup> (Smith 1994), p. 73.



given that they change their moral beliefs.... In other words, the moralist is someone who simply desires, contingently, to do what is right.<sup>14</sup>

I think Smith is essentially making my First Objection.

Here is a simple reply to the First Objection. When the suggestible person discovers that I have listed avocados, she comes to desire avocados in an *original* way, directly, not instrumentally as a way to satisfy her general desire to eat what is on my list. Someone with the *de dicto* desire to eat what is on my list comes to desire avocados only instrumentally, when she learns that I have listed them. So that is an obvious difference. But it suggests the Second Objection.

### Second Objection

The reply just given, saying that the desires that the morally suggestible person is disposed to have are not instrumental desires, contradicts the definition or essence of an instrumental desire. For an instrumental desire is one that is contingent upon a certain belief. That is the only grasp we have on the notion of an instrumental desire.

I don't think Smith has ever made this objection, but it is suggested by the first. The idea is an essentially functionalist one. It is that what makes a desire to  $\phi$  instrumental is that, for some  $\psi$ , it is contingent on a belief that by  $\phi$ -ing one will  $\psi$ . What identifies my desire to be wealthy as an instrumental desire is that it is contingent on my beliefs about what wealth can get me, whereas by contrast someone who has fetishized wealth wants to be wealthy independent of what he thinks money can buy. Since a suggestible person's desire to eat avocados is contingent on her belief that I have listed avocados, that desire is instrumental, and similarly, the morally suggestible person's desire to help the poor is contingent on his belief that helping the poor is right, and so it is an instrumental desire.

But the thought behind this Second Objection seems to be false. For if you are suggestible, and come to believe that I have listed avocados, what happens to you is that you suddenly find yourself wanting to eat avocados. You would find the very thought of eating avocados to drive you wild with desire. Your mouth waters as you contemplate guacamole. Imagine instead a person who would, upon learning that I had listed avocados, feel a very grudging desire to eat some. Her lips would pucker up with disgust, but she would think, "Oh well, I guess I'd better." It seems exactly apt to ascribe to her the *de dicto* desire to eat what is on my list. But it seems particularly *inapt* to ascribe that desire to the suggestible person.

---

<sup>14</sup> (Smith 1997), p. 112.

While this reply seems compelling, it rests heavily on the phenomenology of desire. That is not a good resting place. For the Second Objector might plausibly deny that desires are mainly characterized by their phenomenology, and especially that such a deep and structural feature of a desire as whether it is instrumental could depend on how it feels. I have some sympathy with this claim, myself. Fortunately, I have a more satisfactory, though more complicated reply.

### 3. Two Ways a Desire Can Be Conditional

I said above that a suggestible person's desire to eat avocados is contingent on her belief that by eating them she will eat what is on my list (that is to say, on her belief that avocados are on the list). But there are two ways in which a desire can be contingent. A desire's instrumentality is related to only one of those ways, and not to the way in which the suggestible person's desire is contingent.

So first, a desire can be contingent on some proposition,  $p$  (for example, the proposition that Martha believes I have listed avocados) in the sense that its existence depends on  $p$ . If tomorrow  $p$  is no longer the case, then the desire will no longer exist; if  $p$  had not been the case, the desire would not have existed. A suggestible person's desire for avocados is contingent in this way on the proposition that she believes avocados are listed. But second, a desire can be contingent on  $p$  in the sense that its content, or its satisfaction condition, is itself conditioned on  $p$ . For example, your desire to fly a kite in the park tomorrow is conditional on there not being a hurricane. The desire is not satisfied if you fly a kite in the park tomorrow in a hurricane. It is common to have desires whose existence is contingent on a certain proposition, but whose content is not conditioned on the proposition. For example, Derek Parfit's desire that Venice not sink into the sea<sup>15</sup> is obviously contingent on the proposition that Parfit is alive, for it would not exist if he were not alive and it will not exist when he dies. But its content is not conditional on Parfit's being alive, for its satisfaction conditions in no way entail that he lives. I will say that a desire is *contingent* on  $p$  if its existence depends on  $p$ , and *conditioned* on  $p$  if its content or satisfaction depends on  $p$ .

Now a desire to  $\phi$  that is conditioned on the proposition that by  $\phi$ -ing one will  $\psi$  is, plausibly at least, an instrumental desire; one desires to  $\phi$  only as a means to  $\psi$ -ing.<sup>16</sup> But a desire to  $\phi$  that is contingent on one's believing

---

<sup>15</sup> See (Parfit 1984), p. 151.

<sup>16</sup> Only "plausibly", because there are various complications. For one thing, one might be desiring to  $\phi$  as a way of  $\psi$ -ing, for example, if I desire to run as a way of getting exercise. Furthermore, one might be desiring to  $\phi$ , not as a way or means to  $\phi$ -ing, but as a way or means to doing something that has  $\phi$ -ing as a byproduct, for example, my desire to take an aspirin only if by taking an aspirin I will lower the column of mercury in the thermometer sticking out of my ear. I do not desire to take the aspirin as a means of

that by  $\phi$ -ing one will  $\psi$  is not, at least not necessarily, an instrumental desire. To see this, consider Julius.

Julius has an unpleasant and irrational antipathy to Romans. This antipathy is caused, in some extrarational way, by his belief that one of his ancestors is Roman. Julius does not know which ancestor it is, nor does he have any other feelings or beliefs about this ancestor. It's not that he dislikes the ancestor, is ashamed of him, or anything like that. But if Julius came to believe that he does not after all have a Roman ancestor, he would immediately lose his prejudice against Romans. So Julius desires to avoid Romans, and this desire is contingent on his belief that by avoiding Romans he will avoid people from the birthplace of one of his ancestors. But it is clearly wrong to say that he desires to avoid Romans *in order* to avoid people from the birthplace of one of his ancestors. If this is not clear enough, think of what would happen if Julius came to believe that none of his ancestors was Roman after all. He would lose his prejudice. And then he would certainly not desire to avoid anyone from the birthplace of one of his ancestors. By losing the belief, he would lose not only the contingent desire, but the only plausible candidate for the end toward which the contingent desire might be a means. But instrumental desires don't work that way. If I want to swallow this pill as a means to reducing my fever, and then come to believe that swallowing the pill will *not* reduce my fever, I do not lose the desire to reduce my fever.

So desires that are contingent on one's believing a conditional proposition are not thereby shown to be instrumental desires. The sort of contingency that demonstrates instrumentality is the other kind, the conditionality of a desire. A suggestible person's desire for avocados is not conditioned on the person's belief that avocados are listed, nor is it conditioned on avocados actually being listed. But that is exactly the sort of contingency that is relevant to instrumentality. So the suggestible person's desire for avocados is not instrumental. By contrast, a person who desires *de dicto* to eat what is on my list *does* have a merely instrumental desire to eat what is on my list, for that desire is conditioned on avocados being listed. Similarly, a morally suggestible person's desire to help the needy is contingent on her belief that by helping the needy she will be doing the right thing, but it is not conditioned on the proposition that by helping the needy she will be doing the right thing. Her desire to help the needy is conditioned on nothing at all. So the morally suggestible person desires originally, not instrumentally, to help the needy.

The Second Objection is thus refuted. I don't think it should be surprising that it fails. For remember Williams' point. It was that the man who is

---

lowering the column of mercury, but as a means to something that has lowering the column of mercury as a byproduct. I will ignore these complications.

motivated by the thought that helping his wife is morally right seems somehow defective in his motivations. An agent is a moral fetishist, in our sense, just in case what appeals to the agent about the moral actions that she wants to perform, is that they are moral actions. The good moral agent, by contrast, finds attractive the properties that *ground* the rightness, the right-making properties, or what she takes to be the right making properties. But then it seems very plain that the morally suggestible person is no fetishist. For her disposition is a disposition to desire to  $\phi$  when she comes to believe that  $\phi$ -ing is right. For instance, when she comes to believe that helping the needy is right, she develops a desire to help the needy. And then it is the thought, "By giving money to the Red Cross I will help the needy" that motivates her to give money to the Red Cross. There is no thought of rightness as such in her motivating thought.

I am going to argue that moral suggestibility is not a good Externalist model of moral motivation. First, though, I will introduce another alternative. The problem with the moral suggestibility model will emerge more clearly by comparison with this alternative.

#### 4. Another Externalist Model: Second Order Desires

In reply to the suggestion of Copp's I mentioned above, Smith writes, "Instead of saying, as I suggest, that moralists possess a desire to do the right thing, they [Copp and David Brink] both suggest what makes someone a moralist is the fact that they have a desire to acquire noninstrumental desires to perform acts with right-making features."<sup>17</sup> I don't think that was Copp's suggestion. It may have been Brink's suggestion.<sup>18</sup> But as a matter of fact, whether it was anyone's suggestion at all, it matches nicely an idea introduced by David Schmidtz in another context.

##### 4.1 Second order desires and maieutic ends

Schmidtz argues<sup>19</sup> that even according to a generally instrumental conception of practical reasoning, a person might still be able to choose final ends. A final end is one that is pursued for its own sake and not merely for the sake of something else. Traditionally it has been supposed that an instrumentalist about practical reason cannot accept that final ends might be chosen, or at least not chosen for a reason, since the only practical reasons we can have are instrumental ones. But Schmidtz introduces the idea of a *maieutic* end: "an

---

<sup>17</sup> (Smith 1997), p. 115.

<sup>18</sup> In (Brink 1997). Brink notes, "It is quite plausible that morality itself enjoins intrinsic concern for oneself and one's intimates," (p. 27), and later, "It is not that these forms of special concern causally bring about dutiful action; it is, rather, that they are part of one's duty." (p. 29)

<sup>19</sup> (Schmidtz 1994).

end achieved through the process of coming to have other ends.”<sup>20</sup> As an example, he offers the goal of having a career. You might choose a career in medicine instrumentally, perhaps because you want to have a large income and believe that by having a career in medicine you would have a large income. Or you might just think that careers come with certain goals of their own, goals which are satisfying to pursue. You choose medicine for this reason, an instrumental one. But once you have chosen medicine as a career, the goals you have—of relieving suffering, being a respected member of the medical community, and so forth—are goals you pursue for their own sakes. Having adopted these goals for one reason, you then pursue them not for that reason but for reasons internal to the outlook of your profession.

Notice that some maieutic ends can only be satisfied by the adoption of some fairly specific sort of final end. For instance, you might want to have a career that feels intrinsically rewarding, that “fills your life with meaning,” that gives you a sense of purpose. It is plausible that only a career anchored by final ends like the end of relieving the suffering of the sick can achieve this maieutic end. Becoming the sort of doctor who cares for his patients only instrumentally just won’t do the trick.<sup>21</sup>

A maieutic end is much like a second order desire. To have a maieutic end is to desire to have certain ends, or ends of a certain type. If your end is to have a career, and part of having a career is having certain other ends, then one of your ends is to have ends of a certain kind. Your desire would be to have the desires constitutive of a rewarding career (and also, of course, to have the other things that constitute a rewarding career). You would want to desire to cure the sick, you would want to desire to have the respect of the medical community, and so forth. This want is a second order desire. Now not all second order desires are effective. I might want to stop desiring junk food, but my wanting it doesn’t make it so. But suppose that a particular second order desire, the desire to value for their own sake those things that are (or that one believes to be) morally right, is an effective one. Then someone who had such a desire would be the sort of person Smith thinks that Brink and Copp are describing. And such a person would, I think, be a plausible model for an Externalist of a good moral agent. Let’s call him David. According to this model, “what makes someone a moralist is the fact that they have a desire to acquire noninstrumental desires to perform acts with right-making features” (as Smith puts it in the quotation cited above).

The model passes both tests. David’s motivations will track his changes in moral view, for once he comes to believe that, say, voting for Social Democrats is right, David’s second order desire will kick in and he will

---

<sup>20</sup> (Schmidtz 1994), p. 228.

<sup>21</sup> Thanks to an anonymous referee for suggesting that I spell out explicitly the point in this paragraph.

actually desire to vote for Social Democrats. And David is no fetishist. For his end of having moral ends is maieutic, and the ends to which it gives rise are, once arisen, free standing, final ends rather than instrumental ones. He desires to vote for Social Democrats because (a genuinely causal ‘because’) he believes that it is right to do so, but once the desire is generated it is not conditioned on the belief that voting for Social Democrats is right, any more than our doctor’s desire to cure the sick is conditioned on his belief that medicine would be a satisfying career. The doctor chose medicine *because* he expected it to be satisfying (and lucrative), but once chosen and its constitutive ends adopted, the career is no longer pursued for (merely) instrumental reasons.

#### *4.2 Moral suggestibility is not a second order desire*

Smith is wrong to say that Copp’s model, which is a model of a morally suggestible person, is a model of someone with a second order desire to do what is right. To show that he is wrong, I will provide a new test, a more sophisticated version of the tracking condition. We will see that the moral suggestibility model fails this test, while the second order desire model passes it. The argument will show both that the two models are distinct, and that the second order desire model is superior.

A morally suggestible person meets the tracking condition: her motivations change in step with her moral beliefs. But her motivations do not seem to track her general doxastic state in quite the right way. The example of Ursula shows why not.<sup>22</sup>

Suppose Ursula is a utilitarian. She believes that utilitarian moral theory is true. And suppose that she is morally suggestible. So Ursula desires that the net happiness of sentient beings be maximized. Like many utilitarians, and contrary to unflattering portrayals of utilitarianism, Ursula cares about people (and other animals) and their happiness and suffering; she is not concerned originally with abstract quantities or measures. But although she does believe that utilitarian theory is correct, Ursula has a healthy sense of her own fallibility. She has found herself unable to give answers to challenges raised recently in her moral philosophy class by adherents to rights-based theories. She thinks it is possible, though not likely, that some individual rights should properly act as side constraints against utility maximization.

Ursula is sufficiently self-aware that she knows she is morally suggestible. So she knows that if she ever does come to believe in genuine moral side constraints, she will desire to abide by them. This troubles her,

---

<sup>22</sup> Sarah Wright used an example like my Ursula example, as an objection to the Moral Suggestibility model of moral motivation, in some comments delivered in a seminar at Brown. I thank her for it.

but not in the way such uncertainty might trouble you or me. You or I would, presumably, be troubled by the thought that in our juggernaut pursuit of utilitarianism we might be doing something terribly wrong. What troubles Ursula is that she might someday *correctly* become convinced that a rights theory is correct, and then she would no longer pursue utilitarianism wholeheartedly. Naturally this bothers her a great deal, because it means much of the suffering that she could relieve would remain in the world. For her, there is no compensation in the fact that she would be respecting people's rights, because Ursula doesn't care at all about anyone's rights. This attitude seems strange, but it is the attitude of a morally suggestible person.

It is not the attitude of a person with a second order desire that she desire to do whatever is right—of David, say. David wants to desire to do whatever is in fact the right thing to do. If David thought it possible that moral side constraints were a fundamental feature of true morality, his second order desire would motivate him to investigate. For he wants to have desires of a certain sort (namely, ones whose objects are morally right actions), and he knows that there is a significant chance that the only way he can have desires of that sort is by thorough moral reflection and investigation. So he will be motivated to investigate. But the attitude of a morally suggestible person like Ursula will be more like a fear that she might come to change her mind.

To see that this must be the attitude of the morally suggestible, note that in considering the possibility that there are genuine and important moral rights, Ursula acquires no motivation whatsoever. She does not now care about any such rights, and she has no conditional desire to respect them if they really are well grounded theoretically. Her moral suggestibility provides her with no such concern. The connection between her moral beliefs and her motivations is a pure disposition, not a rational connection. Some analogous examples may help to make Ursula's situation more plain.

Recall Julius, who has the unpleasant and irrational antipathy to Romans, caused by his belief that one of his ancestors is a Roman. But suppose we have caught Julius before he has found out anything about his ancestors. So far, he has no prejudice against Romans, but if he were to discover that one of his ancestors is a Roman, he would come to have that prejudice. Suppose that Julius is keenly self-aware, so that he knows he has this disposition to develop prejudices. Since Julius, like you and me, does not want to be saddled with irrational prejudice, he is glad that he doesn't believe that any of his ancestors is Roman. But he is worried, because he thinks he might just possibly have a Roman ancestor, and he is afraid that if he learns more about his family tree, there is a small chance that he will find some Roman ancestor lurking in his past. Julius's reaction, quite sensibly it seems, is to avoid finding out any more than he already knows about his ancestry. For he does not want to develop a desire to avoid Romans. Of course, he thinks it is unlikely that any of his ancestors are Roman, but there is a chance, and it is

not one he wants to take. Ursula's epistemic situation is analogous to Julius's. That is why she wants to avoid engaging in any more moral theorizing. She is afraid she might possibly learn that individual rights are morally important, and that her suggestibility would be triggered and she would no longer do everything she could to relieve suffering and promote happiness. Unlikely, but not a chance she is willing to take.

Here is a different example. This example will confuse the issue, but only, I hope, temporarily. Mark does not like snails. Mark is suggestible; he is disposed to desire to eat things upon discovering that they are on my list. And he knows that he is suggestible. Though Mark does not believe that I have listed snails, he thinks that I just might do a disgusting, sneaky thing like that, and of course he knows that if he ever does discover that I have listed snails, he will desire to eat them. What should Mark's attitude be toward acquiring further information about my list? An old joke: I don't like snails, and I'm glad I don't, because if I liked them I would eat them, and I can't stand the creepy things. But Mark, no doubt, will at least have no aversion to finding out that I've added snails to my list. He knows he would then desire to eat snails, and he does not now desire to eat snails, but this ought not to *bother* him. Why is Mark's attitude not analogous to Ursula's or Julius's? Is one of the stories wrong?

Even beyond their preferences for foods, most people have a desire to eat things they like. This desire is not a mere 'desire by courtesy', like the general desire that one's desires be satisfied, or the general belief that one's beliefs are true. The desire to eat things you like has independent force and import. For it means that you now desire that you will eat snails five years hence if you will like snails five years hence, even if you do not like them now. And it means that you prefer that *if you did* like snails, you would eat them, even if you do not, as a matter of fact, like them. The desire to eat things you like is, as far as I can tell, just an instance or instrument of the more general desire for pleasurable experiences. Eating things you like is a pleasurable experience. This is why Mark's case is unlike Julius's or Ursula's. For most people, eating snails has no independent importance above and beyond the provision of pleasurable (or unpleasant) experience. It might be different if snails were particularly wholesome or particularly unwholesome. If they were, or if Mark had some other special reason to want to avoid snails, his case would be different.

Ursula and Julius do not expect to gain any particularly pleasurable experience by acquiring and acting on the desires they might acquire if they learned more. Or perhaps they would, but the importance to them of these experiences pales in comparison to other aspects of the desires and their targets. For Julius, the problem is that if he were to discover that one of his ancestors were Roman, and thus acquire a desire to avoid Romans, he would frustrate a desire (or better, a value) that he now has. The prejudice would



interfere with the ordinary important running of his life, and furthermore it is a trait of character which he now finds disturbing. For Ursula, the problem is that if she were to learn that rights really are morally important, and thus acquire a desire to respect them, her present desire to contribute to the welfare of sentient beings would be at least partly frustrated. And this desire, unlike a desire to avoid eating snails, is not conditioned on its own persistence.

Now surely Ursula's is entirely the wrong attitude for a good moral agent. A good moral agent who believes utilitarianism but thinks that there is a chance that a rights-based theory is correct would surely want to investigate the possibility. So while a morally suggestible person does meet the simple tracking condition, in that her motivations follow around her moral beliefs, she does not meet a more sophisticated tracking condition, because her motivations do not respond properly to her moral uncertainty.

I conclude that the moral suggestibility model is distinct from the second order desire model of moral motivation, and that moral suggestibility is not a good model because it fails the sophisticated tracking condition. The second order desire model, on the other hand, looks so far to be an acceptable Externalist alternative.

In the final section, I will consider a challenge to the second order desire model. I will argue that the challenge is not successful, and that the second order desire model is indeed an acceptable Externalist alternative.

### **5. Fetishism at the Second Order of Desire**

As I said, Smith attributes the second order desire model to Brink and Copp. And he objects to this view, he says, for exactly the same reason that he objects to the *de dicto* model of the morally good person, namely, that it is a model of a fetishist, not of a good moral agent. I find his reasoning somewhat difficult to follow, but here is what he says.

It thus is not the case that they [Brink and Copp, as defenders of the second order desire model] are committed to the false view that morally virtuous people are ultimately motivated by the fact that their acts have right-making features, rather, when they act, they are motivated by the features that they believe to be right-making features themselves.

Sure enough, when they act, they are appropriately motivated. But what ultimately moves them, as moralists? All Brink and Copp have managed to do is to reorient the fetish that their so-called morally virtuous people possess. It isn't now about their actions. Rather it is about themselves and their own desires. As I described them they were ultimately motivated by a desire that their acts have right-making features, not be the features that they believed to be right-making features themselves. This seems to me perverse, and Brink and Copp apparently agree. But as Brink and Copp describe morally virtuous people they are ultimately motivated by the fact that they have noninstrumental desires to perform acts with right-making features, not by the fact that they have noninstrumental desires to perform acts with the features that they believe to be the right-making features themselves.... They should therefore agree that the morally virtuous person they describe sounds equally precious, equally self-absorbed, equally fixated on something that isn't of any moral significance: the moral standing of the

contents of his first-order desires, rather than the features in virtue of which his first-order desires have the moral standing they have.<sup>23</sup>

As I understand it, Smith's complaint is that a second order desire with the content, that whatever is right I desire to do, is fetishistic. To clarify, we might again distinguish senses of a desire attribution by disambiguating scope. (Or to put it as Smith did originally, there are two desires, the second order one and its first order object, that might be *de re*.)

(D) D. desires that he desire to do what is right.

This time we have three possibilities.

- (i) Desiree desires that (Desiree desires that ( $x$ ) (if  $x$  is right then Desiree does  $x$ )).
- (ii) David desires that ( $x$ ) (if  $x$  is right then (David desires that David does  $x$ )).
- (iii) ( $x$ ) [If  $x$  is right, then (Dana desires that (Dana desires that Dana does  $x$ ))].

The first attributes to Desiree the desire to be a fetishist.

The third is what Smith thinks must be true for Dana to be a good moral agent. But (iii) could not explain tracking.

The second is the claim that Smith attributes to Brink and Copp, and he thinks it means that David is a fetishist. It's not that David's desire to *act* has been misplaced. It's that his desire to desire things is misplaced. For example, suppose David comes to believe that it is right to end the practice of using chimps in medical research. So he desires that he desire to prevent such use. But David has this second order desire only instrumentally, only as a way of achieving his real goal, which is to desire to do those things which are, in fact, morally right. A good agent would care *originally*, not instrumentally, about desiring to save chimps.

Smith also appears to be suggesting that David's first order motivations are infected with the fetishism of his second order motivations, for he says: "as Brink and Copp describe morally virtuous people they are ultimately motivated by the fact that they have noninstrumental desires to perform acts with right-making features." Whether or not Smith means to say so, it does not appear to be true. There is a sense in which the second order desire is David's "ultimate" motivation, but the ultimacy is only causal. His first order motivations are caused by his second order motivations, but once in place they no longer depend, in any rational way, on the existence or content

---

<sup>23</sup> (Smith 1997), pp. 115–16.

of the second order motive. In Schmitz's terminology, the end of having moral ends is maieutic; once it gives birth to the moral ends themselves, those new ends have their own lives, and might be sustained by other new concerns or merely by their own inertia.

Still, there is a challenge to be met. Even if David is not convicted of *acting* for the wrong reasons, out of the wrong motives, he might be guilty of *desiring* for the wrong reasons.

Does (ii) mean that David is a fetishist? To be moved by the thought of righteousness is to have one's motivations removed from their proper target, which is the characteristics of the action that *make* it right.<sup>24</sup> But now we are wondering whether a similar point holds when you are thinking about what to desire. You are sitting in your office, thinking about Williams' example, and you wonder what would motivate you in such a situation. You hope, let's suppose, that you would be motivated by the thought that it is your wife (or your husband). Why do you hope so? If (ii) is true, David hopes so because he thinks that is the morally best motivation to have. Smith thinks that David's hope ought to be motivated, not by the thought this is the morally right motivation to have, but by the thought that, after all, it would be his wife. That is Smith's complaint against the second order desire model. But Smith is wrong.

First, notice that nobody can complain about the desire itself, the second order desire that (ii) attributes. It is common ground that a person ought to want to be moved by right-making characteristics of actions, and that she ought to want this even if she is not sure what the right-making features of actions are. If she were told that some day in the future she will see clearly what features of actions are the right-making ones, and asked whether she hopes that when that day comes she will be motivated by those features, she surely must say that she does hope so. Otherwise she could hardly be called a good moral agent. So it is not (ii) itself that provides any ground for complaint, not its truth. The ground for complaint would have to be that the second order desire attributed by (ii) is playing the role in the model that really ought to be played by something else.

But second, it would be a mistake to think that the second order desire plays too much of a role. While the second order desire does play an original causal role in generating the admirable first order motivations, it needn't play any maintenance role once the first order motivation is formed. Think of what happens once David has appreciated the rightness of working to end chimp experimentation. He now cares originally, not instrumentally about ending the experiments. And because he cares about ending the experiments, he cares

---

<sup>24</sup> As Korsgaard puts it in (Korsgaard 1986), the reason for acting is not that the action is right; rather, "the reasons why an action is right and the reasons why you do it, are the same." (p. 10)

about those means that are instrumental toward ending them, and in particular, he desires, hopes, to continue to be motivated by the thought of chimpanzee suffering (since he won't act unless he cares, and he wants to act). So now he has another second order desire: he wants to continue to want to end chimp experimentation. This new second order desire is thoroughly *de re*. No thought of rightness enters into it. It is an important desire, I think. A first order desire to end chimp experimentation that was maintained *only* by the general second order desire to want what is right, would seem rather too tenuous, too much like Mark's desire to avoid eating snails.

Third, David's first order desire to end chimp experimentation is not at all conditioned on its being right to do so. This last point may not be obvious. It is an open question, not settled by anything said about David so far, whether his desire to end chimp experimentation is *contingent* on his continuing to believe that it is right to do so. I am not sure, myself, whether a good moral agent's desires would all be contingent in this way. I suspect not. I suspect that if you were persuaded that contributing money to aid the suffering was morally right, that you would continue to care about it even if you were later convinced that rugged individualism was a sounder basis for moral prescription. But in any case, whether contingent or not, the first order desire is not conditioned on the rightness of the action. The simplest way to show this is to examine the formula (ii). The consequent of its quantified conditional is just the formula,

David desires that David does  $x$

and not, in particular,

David desires that David does  $x$  so long as  $x$  is right

The content of the first order desire that is generated by the second order desire is not conditioned on the rightness of doing  $x$ .

Let me sum up.

It would be important, and gratifying, and interesting if the Internalist/Externalist impasse in metaethics could be broken by appeal to Bernard Williams' "one thought too many" point from normative theory. Michael Smith argues that any Externalist model of moral motivation is subject to the fetishism objection derived from Williams' point, or else it fails (what I called) the tracking condition. Smith originally considered only two Externalist models: the *de dicto* desire to do what is right, and the *de re* desire to do what is right. The first is fetishistic. The second fails the tracking condition.

I explained two alternative models that an Externalist might employ: the model of a Morally Suggestible person, and the model of someone with a second order desire that for every right action  $x$ , he desire to do  $x$ . I showed that these models are genuinely distinct, both from the first order *de dicto*

desire model and from each other. And both new models meet Smith's original criteria: neither is a model of a fetishist, both meet the tracking condition. The Moral Suggestibility model, however, fails a more sophisticated tracking condition, since a Morally Suggestible person will not respond in the intuitively right way to her own moral uncertainty. The second model, on the other hand, is an adequate model for an Externalist. It meets even the more sophisticated tracking condition. And, contrary to Smith's complaint, an agent with a second order desire to desire to do what is right is not a fetishist, even at the second order of desire.

Internalists will have to rejoin the dispute with Externalists back at the metaethical level.

### References

- Brink, David O. (1986), "Externalist Moral Realism", *Southern Journal of Philosophy supplement*, pp. 23–42.
- Brink, David O. (1989), *Moral Realism and the Foundations of Ethics*, Cambridge University Press, New York.
- Brink, David O. (1997), "Moral Motivation", *Ethics*, vol. 108(1), pp. 4–32.
- Copp, David (1997), "Belief, Reason, and Motivation: Michael Smith's *The Moral Problem*", *Ethics*, vol. 108(1), pp. 33–54.
- Dennett, Daniel (1980), "The Milk of Human Intentionality", *Behavioral and Brain Sciences*, vol. 3, pp. 428–30.
- Fried, Charles (1970), *An Anatomy of Values*, Harvard University Press, Cambridge, Mass.
- Hare, R. M. (1952), *The Language of Morals*, Clarendon Press, Oxford.
- Horgan, T., and Timmons, M. (1992), "Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived", *Philosophical Papers*, vol. 21, pp. 151–73.
- Korsgaard, C. (1986), "Skepticism about Practical Reason", *Journal of Philosophy*, vol. 83, pp. 5–25.
- Lillehammer, Hallvard (1997), "Smith on Moral Fetishism", *Analysis*, vol. 57(3), pp. 187–95.
- Parfit, Derek (1984), *Reasons and Persons*, Clarendon Press, Oxford.
- Schmidtz, David (1994), "Choosing Ends", *Ethics*, vol. 104(2), pp. 226–51.
- Smart, J.C.C., and Bernard Williams (1973), *Utilitarianism: For and Against*, Cambridge University Press, New York.
- Smith, Michael A. (1994), *The Moral Problem*, Blackwell, Oxford.
- Smith, Michael A. (1997), "In Defense of *The Moral Problem*: A Reply to Brink, Copp, and Sayre-McCord", *Ethics*, vol. 108(1), pp. 84–119.
- Williams, Bernard (1981), "Persons, Character, and Morality", in B. Williams (ed.), *Moral Luck*, Cambridge University Press, New York, pp. 1–19.