

CHAPTER II

Externalism, motivation, and moral knowledge

Sergio Tenenbaum

II.1 INTRODUCTION

For ethical naturalists of a certain stripe, externalism about moral motivation is an attractive option. An influential form of ethical naturalism takes moral properties to be natural properties, while maintaining that moral concepts are not *reducible* to descriptive or non-moral concepts. Moral properties are, on this view, natural properties, and moral terms that refer to these properties play an essential role in the explanation of natural, empirically observable phenomena. These properties are epistemologically accessible to us in the same way that, for instance, physical or chemical properties are accessible to us. According to this view, moral facts are known a posteriori, and the explanatory role of moral terms plays an essential role in moral epistemology; many advocates of such a view will take moral knowledge to be warranted, inter alia, on the basis of “inferences to the best explanation.”¹

On the other hand, an internalist about moral motivation thinks that there is a necessary, conceptual connection between judging that ‘x is morally right’ and being motivated to x.² An ethical naturalist of this kind³ claims, roughly, that moral properties are natural, objective properties and that we learn about their instantiation in the world due to their causal powers. But how could the judgment (or belief) that a certain natural property of this kind is instantiated by a certain action type (or some instances of this type) or a certain consequence of acting in a certain way be *conceptually* connected to a completely different item in our

¹ Although I am ignoring some complications, some advocates of roughly this kind of view include Boyd (1988, 2003); Brink (1989); Miller (1985); Railton (1986).

² This is only one version of ethical externalism. For a long list of possible versions, see Brink (1989).

³ Since this is the only kind of ethical naturalism with which I’ll be concerned in this paper, I’ll just refer to it as “ethical naturalism,” unless the context is misleading, in which case I’ll refer to it as “non-analytical ethical naturalism.”

psychological economy; namely, a motivation to act? Even if one were not inclined to accept the Humean thesis that beliefs cannot motivate on their own in its full generality, the belief or judgment that a certain natural, explanatory property is instantiated by, or is the expected effect of, certain possible actions seems to be the wrong kind of thing to exhibit this kind of conceptual connection to motivation. Thus ethical naturalists tend to accept externalism and explain virtuous and moral behavior by the existence of a desire to act morally, or a desire to perform only actions that are morally right. So moral motivation is explained in terms of a self-standing desire to be moral, which, when coupled with the belief that certain actions are morally right or morally wrong, will engender motivation in a completely unmysterious, Humeanly acceptable way.⁴

However, Michael Smith has mounted an important objection to this form of ethical externalism.⁵ According to Smith, reliance on such a desire to be moral would make the virtuous agent liable to a charge of fetishism; the virtuous agent on this picture would care for those properties of an action that make the action morally right only because, in virtue of having these properties, they make the actions right, rather than caring for the properties that make the action right directly. So rather than being motivated to help because she cares about people in dire need, the virtuous agent in this picture cares for what is morally right, and by realizing that helping people in need is morally right, she cares, derivatively, to help those in need. According to Smith, this gives the virtuous agent “one thought too many,” and does not adequately capture the motivation of such an agent; the virtuous agent should have a *direct* motivation to help those in need. I will argue that Smith’s objection ultimately fails; in particular, I will argue that Smith is wrong to think that the non-analytic naturalist cannot allow that the virtuous agent is typically and reliably motivated by a direct motivation to pursue the specific ends that morality enjoins us to pursue. In responding to this objection, I hope we get a better understanding of how non-analytical naturalism must see the connection between moral knowledge and moral motivation in this picture. My paper does not conclusively show that this account of moral motivation is adequate; in particular, Smith is right that non-analytic naturalism requires that the virtuous agent be directly motivated by a desire to do whatever turns out to be morally right, and for all we know, there is

⁴ For examples of naturalists who accept ethical externalism, see, for instance, Brink (1989) and Railton (1986).

⁵ See Smith (1994).

something problematic about having this motivation.⁶ But if there is a problem in this area, it is *not* the fact that externalism forces the virtuous agent to have “one thought too many.” The externalist has room in her theory to say that the virtuous agent also cares directly for justice, honesty, pain relief, etc. just as well as anyone’s virtuous agent.

II.2 SMITH’S OBJECTION TO EXTERNALISM

According to Smith, an ethical externalist cannot escape being committed to an account of the motivation of the virtuous agent that is incompatible with certain ordinary views. Smith points out that anyone needs to explain why a rational, moral agent reliably changes his or her motivation in accordance with changes in her moral beliefs. Consider a moral, rational agent (Mary), who is persuaded that, contrary to what she used to hold, it is morally right to do X rather than Y when choosing between X and Y. We would expect to encounter a corresponding change in Mary’s motivation; insofar as Mary is virtuous, her change in belief would be accompanied by a corresponding change in motivation. That is, Mary will now be motivated to choose X rather than Y.

According to Smith, the externalist can explain such reliable changes in the virtuous agent’s motivation in accordance with changes in their moral beliefs only in terms of a non-derivative desire to do what is morally right, read *de dicto*. That is, on this picture, we can explain why Mary’s motivational changes follow her changes in moral beliefs only if we think that her motivation to X is derived from her desire to do what is morally right (read *de dicto*). Given that, according to the externalist view, there is no internal connection between moral belief and motivation, it would be an incredible coincidence if every time the moral agent acquires a belief to the effect that X is right, she also acquires a *non-derivative* desire to do X; the externalist can explain such reliable changes only by attributing to the virtuous agent a non-derivative *de dicto*⁷ desire *to do whatever happens to be morally right*.

So if Smith is right, on the externalist’s account, when Mary learns that she ought to provide her children with fun and enjoyment (rather

⁶ In fact, I argue that this kind of motivation is incompatible with some important aspects of moral commitment in “Moral Faith and Moral Reason” (unpublished MS).

⁷ I’ll use sometimes the less precise expressions “*de dicto* desire” and “*de re* desire” for simplicity; but, of course, we are not talking about two kinds of desires, but two different ways of attributing or reporting desires.

than, say, concern herself only with developing their talents⁸), she does not thereby acquire a non-derivative motivation to help the children read *de dicto*; her motivation to help her children is derived from her motivation to do what is morally right and her newly acquired belief that helping children in need is morally right.

Smith's point, of course, shouldn't be understood as a claim that the very existence of a desire to do what is morally right, read *de dicto*, is objectionable. On the classic way to understand the distinction in belief contexts, someone holds a *de dicto* belief, at least typically, if the believer would assent to the proposition in the that-clause.⁹ Extending this point to desire, we would expect that to desire something *de dicto* would also be implied, typically, by the fact that the agent would want the object of the desire as described in the content of the desire ascription.¹⁰ Let us suppose that Larry has a daughter Mary who also happens to be the captain of St. Mary's high school football team, and who is in the running for the MVP award for the 2010 football season. Larry, as a good father, wants Mary to do well. Let us now compare the following sentences:

- (1) Larry wants Mary to get the MVP award.
- (2) Larry wants his daughter to get the MVP award.
- (3) Larry wants the captain of the St. Mary's high school football team to get the MVP award.

If Larry knows that Mary is his daughter and that she is the captain of the high school football team, in a very clear sense all of (1)–(3) are true of Larry even on a *de dicto* reading; these are all things that he wants and he is disposed to pursue the objects of these desires under any of these descriptions. So, for instance, if he reads that the captain of a high school team can win an MVP award only if she has a signed form from her principal, he will be disposed to make sure that the principal signs the relevant form. Similarly, if Larry comes to believe that Mary wins the award, all the following will be true of her even when read *de dicto*:

- (1) Larry believes that Mary got the MVP award.

⁸ A view inspired by Roger Scruton's comment about his bringing up his son: "It goes without saying that Sam will not enjoy his childhood ... But that is not the point. Childhood is not an end in itself but a means to growing up." Cited in Heller (1999).

⁹ Obviously there are other ways to understand the distinction. There is also an understanding of "*de re* beliefs" as a relation to an actual object, or as a difference of the scope of the existential quantified relative to the attitude verb, but these readings are obviously even further removed from the sense that Smith intends in the discussion of the fetishism objection.

¹⁰ I am using this more convoluted formulation to avoid commitment to the view that the desires and wants are attitudes whose content is necessarily propositions or sentences.

- (2) Larry believes that Larry's daughter got the MVP award.
- (3) Larry believes that the captain of St. Mary's high school football got the MVP award.

It follows that on any account the virtuous agent has a *de dicto* desire to do what is morally right. For, presumably, the virtuous agent believes (read *de dicto*) that it is right to do what is morally right. Thus it is a consequence of Smith's practicality requirement that if the agent now does not suffer from any relevant form of irrationality, as is the case of the virtuous agent, she will also be motivated to do what is morally right.¹¹ But now there seems to be no reason to think that she would not recognize this as a description of what she wants. In other words, we can ascribe a desire to be moral to a virtuous agent even in opaque contexts without making the agent less virtuous for this reason.¹² Moreover, on everyone's view, the virtuous agent has a *de dicto* desire to pursue X, once she recognizes that X is morally right. On both views, once the virtuous agent *recognizes* that, say, helping the poor is morally right, then, she has thereby a *de dicto* desire to help the poor, at least because she recognizes "helping the poor" as a correct description of what (or of one of the things) she desires. So we cannot find any difference regarding *de dicto* and *de re* desires when it comes to Smith's views or the externalist's with respect to the virtuous agent. Moreover, given that, as I'll argue shortly, the point seems to be about the *basing* or the *inferential status* of one's desire, specifying that the desire is *de dicto* is unnecessary; after all, the contexts in which we are examining the agent's inferences or the basing of her beliefs are typically contexts in which the attitudes are being ascribed *de dicto*.

Even though all views in question attribute the same desires to the virtuous agents, they don't attribute it in the same way. According to Smith, the externalist has trouble explaining how the virtuous agent could have

¹¹ More precisely, the practicality requirement says: "If an agent judges that it is right for her to Φ in circumstances C, then either she is motivated to Φ in C or she is practically irrational" (Smith 1994). But since the antecedent is presumably true of the virtuous agent when we replace " Φ " with "act morally" or "do that which is morally right" and "C" with any circumstance, it follows that the virtuous agent has a *de dicto* desire to do what is morally right.

¹² Jonas Olson, for instance, argues that a virtuous agent could have both kinds of desires (Olson 2002). If I am right, Smith is already committed to accepting that the virtuous agent has both kinds of desires. But I do not see why he needs to deny this; the point is that only the virtuous agent can reliably acquire *de re non-derivative* desires. Olson also points out that often moral action can be motivated only by a *de dicto* desire to do what is morally right in a non-objectionable way. This is probably correct, but all that Smith needs to make his point is that the virtuous agent needs to be reliably motivated by a non-derivative *de dicto* desire to do X when X is morally right, not that she must be *always* so motivated.

certain *non-derivative* motivations that are obviously included in our ordinary understanding of the motivational makeup of a virtuous agent:

Good people care non-derivatively about honesty, the weal and the woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right, when this is read *de dicto* and not *de re*. (Smith 1994: 75)

But the question is how to spell out the notion of “derivative” and “non-derivative” that gives rise to a serious difficulty for the externalist, or at least for our kind of non-analytic naturalist externalist. As we’ll see in the next section, this is not as easy as it might seem.

II.3 DERIVATIVE, NON-DERIVATIVE, AND INSTRUMENTAL MOTIVATION

We can start by assuming that “derivative” refers to some kind of rational or inferential connection (or a causal connection that obtains in virtue of, or is constitutive of, the agent’s rationality); it concerns the basing of a desire, or how it is inferentially arrived at. After all, the fact that a certain motivation just has a certain causal genesis that involves other motivational attitudes shouldn’t be, at least not systematically, enough to disqualify it from being the right kind of motivation. So, if Larry fell in love with his wife because he was attracted to chess players, or because they met at a time when he was emotionally vulnerable and thus disposed to engage in charged conversations with recent acquaintances, then these motivations are in the causal path to his motivation to help his wife whom he now loves. But the fact that these items figure in the causal genesis of Larry’s motivation to help his wife does not make his desire to save her life “derivative” in any objectionable way.¹³ On the other hand, an agent’s motivation to save his wife might be objectionable if the agent cares for his wife only insofar as morality demands that he cares for his wife; this kind of motivation seems objectionable exactly because of the way in which the agent’s desire to help his wife is *inferentially arrived at*,¹⁴

¹³ Of course, there might be cases in which the causal genesis of a desire might make the motivation problematic in various ways. However, it would be hard to claim that it makes it “derivative” in this way.

¹⁴ “Inferentially arrived at” here needs to be understood in a weak sense; we do not want to say that the desire is objectionable only if it is arrived at by explicit deliberation. I’ll leave it open how to understand “inferentially arrived at” more precisely; all that my argument requires is that not all causal paths leading to a desire count as inferential paths.

or based on,¹⁵ the desire to perform only actions that are morally right. We can think about two “extremes” about what would make a certain desire or motivation derivative:

[MAXIMUM] A motivation to Φ is derivative if it is inferentially arrived at from (or based upon) any other mental states.

and:

[MINIMUM] A motivation to Φ is derivative¹⁶ only if it is inferentially arrived at from (or based upon) a belief that Φ -ing is causally necessary to bring about that p together with a desire that p .

In a nutshell, MAXIMUM counts as derivative all but what Nagel would call “unmotivated desires,”¹⁷ while MINIMUM counts as derivative at most one’s instrumental desires.¹⁸ But we can see that even given the setup of the objection to externalism, one is already committed to rejecting MAXIMUM. After all, in our example Mary is *persuaded* that it is morally right to provide her children with fun and enjoyment. But now it seems that the beliefs that she formed in arriving at the conclusion that it is morally right to provide her children with fun and enjoyment *are* part of the basing of her motivation to provide her children with fun and enjoyment. In fact, [MAXIMUM] makes her motivation inappropriately derivative even when her motivation to do so is simply based on her normative belief that she should provide her children with fun and enjoyment; arguably, even Smith himself is committed to this view. This would make not only the externalist account of moral motivation defective, but also any account that took moral motivation to be typically based on normative beliefs. Obviously, this does not show that the objection fails, but that [MAXIMUM] is far too strong a condition. On the other hand, [MINIMUM] would deliver Smith’s argument only if the naturalist were committed to the view that, for instance, “helping the poor” is a causal means to “doing what is morally right.” But the externalist

¹⁵ That is, the desire to do what is morally right is part of the *reason* or the *basing* of the desire to save his wife. Here, too, the question of how to understand the basing relation is a rather controversial one, but, again, all that my argument requires is that the basing relation cannot be understood as just.

¹⁶ Of course, we are defining “derivative” here with the assumption that derivative in this sense is a form a derivative motivation that would be in some way inappropriate for a moral agent; that is, the kind of derivative motivation to do those things that are morally right that we would not expect to find in a virtuous agent.

¹⁷ See Nagel (1970).

¹⁸ David Brink at some points seems to claim that the fetishism objection would apply only to instrumental desires. See Brink (1997).

need not say that the relation between “helping the poor” and “being a morally right action” is of this kind; the externalist need not say that by performing an action that instantiates the property “helping the poor” one has brought about, as a separate effect, the instantiation of another property; namely, “being a morally right action.” The more plausible view is that “helping the poor” is just a way of, or an instance of, “being a morally right action.” No externalist is committed to the view that helping the poor is an instrumental means to being a morally right action, and the non-analytic naturalist is particularly unlikely to make this move; the naturalist will probably identify the property of being morally right with a complex property that will have things like “helping the poor” among its constituents.¹⁹

One obvious intermediate option is to restrict the class of derivative motivations to motivations inferentially arrived at from, or based on, *other desires or motivational states*. We can define this intermediate position as follows:

[INTERMEDIATE 1] A motivation to Φ is derivative if it is inferentially arrived at from, or based on, a motivational state to Ψ (or that p) and certain beliefs that connect²⁰ Φ -ing with Ψ -ing (or p).

So this would still make the externalist version of the virtuous agent a case of inappropriately derivative motivation. However, it is difficult to accept [INTERMEDIATE 1]. Let us look at one of Smith’s rational requirements of motivation, the requirement to acquire more general desires that provide systematic justification for our existing particular desires.²¹ So insofar as an agent acquires new motivation in virtue of the realization that adding a new desire to her motivational set would give her a more coherent motivational set, her new desire would be derivative under [INTERMEDIATE 1]. But suppose Shoshana improves her moral outlook exactly by making her motivation more coherent; suppose she notices that she treats teachers inexplicably differently from the way she treats lawyers, and, as a result of this realization, and independent of any moral

¹⁹ In *The Moral Problem* (1994), Smith carefully uses the terms “derivative” and “non-derivative” to describe what the externalist is committed to. In a later paper, Smith uses “instrumental” and “non-instrumental,” which does suggest that he takes “derivative” in the sense described by [MINIMUM]. See Smith (2004c). But, again, the externalist need not be committed to the claim that the more specific desire is derivative in the sense defined by [MINIMUM].

²⁰ This has to be left vague, since the whole point of this condition is to rule out *any* inferential move from one kind of motivation to another.

²¹ Smith (1994: 159–61). Smith explicitly classifies at least the more general desires arrived at in this way as “underivative” (159–61).

argument, she changes the way she treats teachers. It seems that Shoshana is just as virtuous as if she had changed the way she treated teachers by an explicit argument to the effect that her differential treatment of lawyers and teachers was morally wrong. But [INTERMEDIATE 1] would classify Shoshana's new motivation as derivative. Another obvious candidate is to extend the category of means in [MINIMUM] to include all constitutive means and make it a sufficient, rather than necessary, condition for a motivation being derivative:

[INTERMEDIATE 2] A motivation to Φ is derivative if it is inferentially arrived at from, or based on, a belief that Φ -ing is necessary (causally or constitutively) to bring about that p (or to ψ) together with a desire that p or to ψ .

The first problem with [INTERMEDIATE 2] is that it is difficult to see why some typical cases of constitutive means make the motivation derivative. For instance, suppose you are spending the afternoon with your daughter and she asks: "Why did you take the time off work to go for a walk with me in the park." Suppose you say: "because I want to spend more time with you," and your daughter says, disappointed: "I thought you just wanted to go for a walk with me in the park." There is a way of interpreting your answer that makes your daughter's disappointment justified. If you had this grand plan of being a better parent, and you realize that being a good parent involves among other things, spending more time with your daughter, and you realize that today there is a good opportunity to put the plan into action and thus take your daughter for a walk in the park, while at the same time you feel jealous about your colleagues who are at this moment in a meeting discussing next year's departmental budget, there is no doubt that your daughter could, and perhaps should, be expecting more from you. But if you simply enjoy spending time with your daughter, and you notice that you'd like to do more of this, and one of the activities that would count as "spending time with your daughter" that you particularly enjoy is taking walks with her away from the city, and you feel particularly eager do this today and in the park, then your daughter probably could not and certainly should not expect more from you. Smith would like to classify the externalist type of motivation as falling into the first category, but [INTERMEDIATE 2] does not discriminate between the two. In fact, it is not clear that the difference between the two can be cast in terms of one motivation being more derivative than the other, or derivative in a different way. "Helping the poor" can be constitutive of the end of "helping the needy" or of the end of "expressing one's charitable disposition." Arguably, the latter is a

case of “one thought too many” but not the former.²² Defining “derivative” in such a broad manner cannot distinguish between these two kinds of motivations.

It was worth looking at another, rather obvious, problem with [INTERMEDIATE 2]. At the limit, one of the things that is constitutive of helping the poor is helping the poor itself; and thus [INTERMEDIATE 2] threatens to trivialize the distinction between “derivative” and “not derivative.” Of course, it is not hard to think of various ways one can try to fix this problem, but it is not clear that we can avoid trivialization without jeopardizing the criticism against at least the naturalist externalist. For, after all, the non-analytic naturalist typically claims that the property of being morally right is identical with a certain natural property. But if this natural property is the one that we should want to instantiate for its own sake, then the naturalist would not be guilty of making moral motivation derivative in any objectionable way. So, for instance, if the only property that virtuous agents are responding directly to is the maximization of general expected utility then by responding to the property of being morally right, a virtuous agent is responding to the property of maximizing expected utility.

But this might lead us to a more perspicuous way of putting the potential problem for the externalist in this vicinity. One might argue that morality calls for acting on motivations whose content is, at least potentially or for all the agent knows, distinct from what morality actually requires. But the externalist cannot explain how a virtuous agent can have her motivation reliably track her moral views and at the same time allow that the virtuous agent will, often if not always, be motivated by desires whose objects are indifferent to changes in her moral views. In fact, in answer to objections, Smith says:

But if the only mechanism that exists for causing and sustaining moral motivation ... is the desire to do what is morally required, then it seems ... that they [sc. morally virtuous people] will positively eschew any non-instrumental concern [for their family and friends]. After all, a non-instrumental desire to look after family and friends wouldn't be kept in check by a desire to do the right things under conditions of moral belief revision. (Smith 2004c: 286)

²² This means that “one thought too many” might be a misleading way to express the problem. One might think that since Smith restricts his claim to changes of mind regarding “fundamental values,” these issues are irrelevant. But this is not correct; fundamental values can be constitutive of other fundamental values. Note that Smith's own list of things that the virtuous agent cares non-derivatively about (cited above) includes fundamental values that seem to stand in this relation (e.g., “people getting what they deserve” and “justice”).

One might think that this argument sidesteps the need for coming up with a more precise understanding of “non-derivative” or “direct” motivation. For we already know that the motivation attributed by the externalist cannot be of the right kind, since in the externalist picture, the virtuous agent can be reliably motivated to do what is morally right only if she *eschews* the right kind of motivation. But this kind of criticism underestimates the resources available at least to the naturalist version of externalism, or so I’ll argue.

11.4 A REPLY TO SMITH’S ARGUMENT

Let us use a toy naturalist theory in which “morally right” is identified with the complex property that is instantiated by any possible action which gets at least a certain minimum score in a weighted sum of various graded properties. So, for instance, “causing pain relief to degree *d*” might be one of these properties, and “instantiates helping a friend to degree *d*” might be another one, and, of course, some properties might have negative weights (“causing suffering”, etc.);²³ the weighted sum also settles disputes about what is morally better. For simplicity’s sake, I’ll assume that, on this theory, there is always one morally right action, the one that is morally best.²⁴ Although I will focus on this toy theory, I hope that it’ll be clear that various (and more plausible) versions of naturalism could use the same kind of response to Smith’s objection.²⁵ Now suppose an action is morally right in virtue of being a case of helping a friend to a very high degree (and to make things easier let us suppose that there is no other available action that is also morally right). Now the externalist is not committed to the view that the virtuous agent needs to figure out all the weighted sums in order to be in a position to know that this is the morally right action. In fact, it is compatible with the externalist that an agent knows which actions are morally right by being able to reliably tell, for instance, when actions exemplify enough of one of these properties. So in our case the virtuous agent might know that:

- (A) This action is morally right in virtue of instantiating property X to degree *d*.

²³ Of course the second one is not a clear example of a natural property. But for our purposes it doesn’t matter.

²⁴ I’ll ignore ties, too.

²⁵ So, for instance, Richard Boyd’s view that, roughly, identifies the property of moral goodness with a cluster of properties could easily adapt the account of moral motivation expounded here. See Boyd (1988).

Moreover, it is compatible with the externalist view that the virtuous agent also cares, at least in the ideal case, for each of the component properties, and thus that she would have an independent motivation to relieve pain or help a friend. So in this case the agent knows that a certain action is morally right, is motivated to do what morality requires, but the *de dicto* desire to do what is morally right plays no essential role in her motivation. Obviously this does not yet answer Smith's concern. It is unclear how a virtuous agent, on this picture, could change her mind about what is morally right and reliably act accordingly; after all, her motivation in this picture does not seem to dovetail changes in moral judgments. But here the desire to do what is morally right can perform a different, but also completely unobjectionable role. It is easier to see this case if we look first at fundamental changes in moral beliefs regarding the relative importance of different constituent properties; in our toy theory, this would amount to changes in beliefs regarding relative weights of the component properties. But since, *ex hypothesi*, the virtuous agent already has the direct motivation to act so as to bring about the instantiation of either property, all that changes now is that the desire to be moral adds motivation to give more weight to one of these component properties.

For instance, suppose that Mary used to think that one should always tell the truth even when it might hurt someone's feelings. But now, after talking to her friends, she's convinced that sometimes she should lie to protect her friend's feelings; in particular, she's now convinced that she should not tell Larry that his poetry is juvenile. Before having these conversations, Mary was ready to visit Larry and let him know that his poetry was juvenile, but now she instead tells him that she finds his poetry very deep and inspiring.²⁶ Before and after her conversation Mary cared directly about honesty and the weal and woe of her friend. However, the desire to be moral first required that she gave more weight to honesty and later to not to hurt her friend's feelings, and the extra motivational force of the desire to be moral tips the balance in favor of either desire, depending on her belief about what is morally right. But it is not clear what could be objectionable about the motivation to be honest playing *that* role.

We can see this more clearly if we think that there are two independent desires that correspond to two different types of moral beliefs. One kind of moral belief is about which properties are components of the "morally

²⁶ Perhaps needless to say, I am not endorsing the moral judgment here.

right,²⁷ and a corresponding desire to engage in actions that are morally right (that exemplify such properties). Now these beliefs generate the *direct* motivation to be honest and to care about one's friends' feelings. This motivation counts as direct, because the virtuous agent does not believe that such things are *instrumental* in bringing about what is morally right, but they are *instances* of the morally right. The second kind of moral belief is a belief about the various weights that are attached to the graded properties that constitute what is morally right;²⁸ in our case, these would be the newly acquired belief that it is more important not to hurt one's friend's feelings. And here we can think that the moral agent has a desire to do what is morally *best*, a desire that earlier "sided" with the direct motivation to be honest and now with the direct motivation not to hurt her friend's feelings.

It is hard to see now why this latter kind of motivation would be objectionable. After all if asked: "Why did you lie to your friend; don't you care about honesty and truth-telling?" Mary's answer could appropriately be something like: "I care about both, but in this case, it was more important not to hurt my friend's feelings." But this fits well with what my externalist naturalist would say about this case; it seems plausible to think that this kind of statement expresses that one's motivation was settled by the desire to do what is right when those two competing motives were in play. Accepting this point does not imply that the virtuous agent needs to eschew the direct motivation to look after friends or be honest, given that this kind of direct motivation plays an essential role in bringing about his virtuous actions.

It is worth noting that one could easily modify this naturalist response to Smith's objection even if one thought that the motivation that is derived from this kind of belief about what instantiates moral rightness is also objectionable. For as long as our naturalist does not have too long a list of component properties, it can be part of the conception of a virtuous agent that the virtuous agent must have an independent direct motivation to instantiate, or bring about the instantiation of such a property. So under this conception part of what makes someone count as a virtuous agent is that she has a direct motivation to be honest, to care for the well-being of her friends, etc. The rest of the solution remains the same, except now the *only* role the desire to be moral has is to side with the motivation

²⁷ Something roughly along the lines of "which properties are *pro tanto* good."

²⁸ Roughly, these would be beliefs about how to move from judgments of *pro tanto* good to judgments about overall good.

that, according to the agent's moral beliefs, is the one that follows that balance of moral reasons.²⁹

Of course, this modified solution seems to fail if the change in virtuous agents' belief is of a more radical nature. The modified solution does not seem to apply to the case in which the moral agent comes to believe for the first time that an action can be morally right in virtue of instantiating a property that he previously thought was morally irrelevant. Suppose Jane is convinced for the first time that loyalty to one's country is morally right, something that should count in favor of our actions. Suppose she had not previously had any direct motivation to perform patriotic acts. On the account we have, this new belief about what is morally right cannot guarantee that, insofar as she is rational, Jane will now have a direct motivation to be loyal to her country. But it is not clear that it is part of the ordinary conception of morality that change in belief could produce direct motivation in this way. Here it might be the case that Jane can only perform her duty to her country from her desire to do what is morally right.³⁰ It might be the case that a virtuous person would, in such a situation, have a second-order desire to acquire the relevant direct motivation, but it doesn't seem so intuitive that we need to think that a virtuous agent must, just in virtue of her rationality, care directly for patriotic acts as soon as she is persuaded that it is morally right to perform patriotic acts, even if she previously saw *no* value in them and had *no* motivation to perform patriotic acts.

²⁹ Perhaps one of the moral beliefs of the moral agent is that she must acquire this kind of direct motivation, and thus the moral agent would have a second-order desire to directly desire certain things. On this view, one does not count as a virtuous agent until this desire is at least partly satisfied. Jamie Dreier also suggests that second-order desires can help answer Smith's challenge to the externalist. However, the second-order motivation to acquire direct motivation to X whenever one forms the belief that X is morally right cannot, on its own, make it the case that the virtuous agent reliably acquires the direct motivation when she changes her moral beliefs insofar as she is rational. For rationality alone does not guarantee that we have the first-order motivation that we desire to have. So, for instance, take Frankfurt's famous case of the psychiatrist who wants to (intrinsicly) desire taking drugs so that he'll know how it feels to be a drug addict. The second-order desire, on its own, cannot guarantee that the psychiatrist actually desires (intrinsicly) to take drugs, no matter how rational the psychiatrist is. Similarly my desire to intrinsicly desire to eat Brussels sprouts (rather than just desiring it as means to better health) unfortunately does not get satisfied just by virtue of my being rational. All that Dreier can show is that the moral agent on this picture will *want* to acquire the first-order motivation, not that he'll acquire it. Of course, one could simply postulate a mechanism in the virtuous agent that guarantees the formation of the first-order desire. But then one might as well postulate a mechanism that works directly from the moral belief (Dreier 2000).

³⁰ This parallels some of the cases that Olson considers (2002). See also Hume's interesting discussion of how the sense of duty can be a secondary motive to perform virtuous actions (2000 [1739]).

Smith is right that the naturalist can explain the reliable motivation of a virtuous agent only by postulating a non-derivative desire to do what is morally right, and, ultimately, there might be something problematic about thinking that such a desire plays a prominent role in moral motivation. Is it really plausible to think that a virtuous agent cares directly about morality as such, independently about the particular ends that she judges to be morally right. Perhaps other aspects of non-analytic naturalism make this picture of moral motivation even more problematic. It seems strange to think that virtuous motivation would change its direction, for instance, simply by a new discovery that certain properties play a certain role in the explanation of social facts.³¹ Perhaps this kind of moral motivation is incompatible with other features of moral commitment. I have not tried here to allay all these concerns.³² But whatever problems there might be with the *content* of a direct motivation to do what is morally right, the existence of such a motivation, or even the role it has to play in the externalist account, is not incompatible with the direct motivation to pursue specific ends that, as Smith rightly observes, must be part of the motivational makeup of the virtuous agent.

³¹ Adams (1999) raises an objection along similar lines; Boyd (2003) tries to provide an answer Adams' objection.

³² In fact, in "Moral Faith and Moral Reason" (unpublished MS), I argue that the non-analytic naturalist version of externalism is incompatible with important features of moral commitment.