

Caj Strandberg
University of Oslo
c.s.strandberg@ifikk.uio.no

Original Scientific Paper
UDC 17.021.1/.2
17.034:165.6
17 СМТ М.

Received: August 28, 2024

Revised: October 01, 2024

Accepted: October 04, 2024



SMITH ON THE PRACTICALITY AND OBJECTIVITY OF MORAL JUDGMENTS

Abstract

The moral problem presented by Michael Smith in his seminal book with the same name consists of three claims that are intuitively plausible when considered separately, but seem incompatible when combined: moral judgments express beliefs about objective moral facts, moral judgments are practical in being motivational, and beliefs are unable to motivate by themselves. An essential aspect of Smith's solution to the moral problem is the contention that moral judgments are both motivating for rational agents and objective. In this paper, I take a close look at Smith's arguments by considering his characterizations of rationality understood as coherence between attitudes. It is suggested that on this understanding of rationality and coherence, it has not been clearly shown that moral judgments are both practical and objective.

Keywords: Michael Smith · the moral problem · moral objectivity · practical rationality · motivational internalism · coherence

1. Introduction

The publication of Michael Smith's *The Moral Problem* in 1994 was a main event in metaethics. At that time, I was a fresh student with a burgeoning interest in metaethics and I can still recall the fascination and admiration to which the book gave rise among teachers and fellow students. Our response was confirmed by the deep impact of the book. The influence of Smith's work is evinced by other philosophers discussing issues it brings to attention and arguments it presents. What is more, it is confirmed by the numerous critical comments his work generated and still generates. As often is the case in philosophy, the better a work is, the more it gets criticized. Further discussion of Smith's work is thus further evidence of the importance of his contribution.

In some respects, the metaethical debate looks different than it did in 1994. One key difference concerns how philosophers conceive of the relation between normative reasons and practical rationality. According to the approach represented by Smith, normative reasons should be explained in

terms of rationality.¹ In brief, an agent has a normative reason to perform an action insofar as she would have a pro-attitude of some sort towards the action if she were practically rational. In recent times, it has been common to assume, conversely, that rationality should be explained in terms of responding to reasons.² In brief, an agent is practically rational insofar as she responds to the normative reasons of which she is aware.³ This is not the place to evaluate the *pros* and *cons* of respective strategy. However, in one important respect I think the former approach has an important advantage: It has the potential of being more explanatory potent. On this strategy, it might be possible to explain reasons in terms of rationality and then explain rationality in terms of coherence between attitudes. Thus, one normative notion—reasons—is explained in terms of another normative notion—rationality—of which one provides a substantive account by employing the notion of coherence. By contrast, on the latter strategy rationality is explained in terms of normative reasons of which it seems difficult to say anything relevantly informative.⁴ Thus, one normative notion—rationality—is explained in terms of a more fundamental normative notion—reasons—and not much substantive can be said about the latter. The present paper can be seen as a discussion of whether Smith's particular version of the former strategy turns out to be successful.

2. The Moral Problem

The moral problem consists in the fact that three claims that, considered separately, are intuitively plausible seem incompatible when combined with one another.

The first claim states that moral judgments are objective:

(1) *Objectivity of Moral Judgments*: 'Moral judgements of the form "It is [morally] right that I ϕ " express a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.'⁵

According to Smith's view of moral objectivity, (1) entails that (i) moral judgments consist in beliefs that have truth value and (ii) there are non-arbitrary criteria of what count as relevant considerations of whether an

1 See e.g. Williams (1981), pp. 101–113; Korsgaard (1996), pp. 188–221, and Markovits (2014), Ch. 3. In Strandberg (2024), pp 256–273, I put forward a Neo-Humean view that represents a version of this approach. In what follows, 'rationality' refers to practical rationality.

2 See e.g. Parfit (2011), Ch. 1; Kiesewetter (2017), Ch. 7, and Lord (2018), Ch. 1–2.

3 The shift in the debate has been motivated by considerations as regards the connections between reasons, rationality, and different interpretations of rational requirements. See e.g. Broome (2013), Ch. 7–10 and Kolodny (2005), pp. 509–563.

4 However, see e.g. Gert (2004), Ch. 7 and Schroeder (2007), Ch. 4.

5 Smith (1994), p. 12. Concise accounts of Smith's view are found in Smith (1996a), pp. 277–302 and Smith (1998), pp. 149–172. See also Smith (1989), pp. 89–174.

action is right, where these criteria are set solely by the circumstances of situations. As Smith puts it, 'the only relevant determinant of the rightness of an act is the circumstances in which the action takes place. If agents in the same circumstances act in the same way then either they both act rightly or they both act wrongly'.⁶

The second claim states that there is a conceptually necessary connection between moral judgments and motivation, a view usually referred to as 'motivational internalism':⁷

(2) *Practicality of Moral Judgments*: 'If someone judges that it is [morally] right that she ϕ s then, *ceteris paribus*, she is motivated to ϕ '.⁸

However, Smith recognizes that an agent might judge that an action is right for her to perform but still not be motivated to perform it if she suffers from some form of practical irrationality, such as weakness of will, depression, apathy, addiction or compulsion. He therefore thinks that (2) should be reformulated in the following way:

'If an agent judges that it is [morally] right for her to ϕ in circumstances C, then either she is motivated to ϕ in C or she is *practically irrational*'.⁹

The third claim states a standard view of motivation:

(3) *Humean Theory of Motivation*: 'An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences'.¹⁰

These three claims create the moral problem: According to (1), moral judgments express beliefs about objective moral facts. According to (2), moral judgments are necessarily accompanied by motivation in rational agents. According to (3), beliefs and desires are distinct existences such that beliefs cannot motivate to action but only desires can do so. Thus, (1) and (2) entail that there is a necessary connection between beliefs about objective moral facts and desires. However, according to (3) this cannot be the case.

6 Smith (1994), p. 5.

7 For a discussion of different types of motivational internalism, see Björklund et al. (2012), pp. 123–137. In Strandberg (2011), pp. 341–369 and Strandberg (2012), pp. 87–122, I propose a version of externalism that employs the notion of generalized conversational implicature to explain the close connection between moral utterances and motivation. In Strandberg (2013), pp. 25–51, I argue that Smith's version of internalism faces a dilemma that can be solved by this version of externalism. Smith's main argument against externalism and for internalism is the much discussed fetishist argument (Smith (1994), pp. 71–76 and Smith (1997), pp. 11–117). In Strandberg (2007), pp. 249–260, I defend externalism from this argument and provide an explanation of moral motivation.

8 Smith (1994), p. 12.

9 Smith (1994), p. 61. Emphasis added.

10 Smith (1994), p. 12. For a critical discussion of the Humean theory of motivation, see Arruda (2019), pp. 157–178.

As can be seen, Smith talks about *the* moral problem. However, strictly speaking, I think there are *two* moral problems. The first concerns the connection between (1) and (2). The problem is that if (1) moral judgments consist in beliefs about objective moral facts, it is difficult to see how (2) moral judgments could be necessarily accompanied by motivation. The second concerns the relation between (1), (2), and (3). The problem is that if (1) moral judgments consist in beliefs and (2) moral judgments are necessarily accompanied by motivation, it follows that there is a necessary connection between beliefs and desires, which seems incompatible with (3). In order not to complicate matters, I will not distinguish between the two moral problems in what follows.

3. A Simple Solution to the Moral Problem?

It might be suggested that there is a simple manner to solve the moral problem. Consider the first characterization of (2), the Practicality of Moral Judgments, according to which it is conceptually necessary that if an agent judges that it is right that she ϕ s, then she is, *ceteris paribus*, motivated to ϕ . That is, on (2) moral judgments are by conceptual necessity *accompanied* by motivation. Now, this can be explained in different manners. According to one alternative, (2) is explained by it being conceptually necessary that a moral judgment partly or wholly *consists* in a motivational state. In the literature, this view is often formulated by saying that motivation is ‘internal’ or ‘intrinsic’ to moral judgments. According to another alternative, (2) is explained by the fact that we *classify* an agent’s judgment as a *moral* judgment only if she is accordingly motivated. This alternative is compatible with the nature of the judgment itself not playing any part in the explanation of why the agent is motivated. In the literature, it is standardly presumed that (2) should be understood in accordance with the first alternative, and it is only fairly recently that it has been noticed that it is compatible with the second alternative.¹¹

We might now return to the moral problem. The problem is assumed to consist in the following: (1) and (2) entail that there is a necessary connection between objective moral beliefs and desires, but according to (3) this cannot be the case since beliefs are unable to motivate and only desires have this capacity. However, this contention presumes that (2) is read in accordance with the first alternative. It is only if (2) is understood to entail that a moral judgment wholly or partly consists in a motivational state that it together with (3) entails that moral judgments cannot consist in beliefs. However, if (2) is read in accordance with the second alternative, (1), (2), and (3) are fully compatible. According to this alternative, we classify a judgment as a moral judgment only if it is accompanied by motivation in the form of desires, which is fully compatible with moral judgments consisting in beliefs about objective moral facts.

11 See Tresan (2006), pp. 143–165. See also e.g. Strandberg (2011), pp. 342–347 and Francén (2020), pp. 366–379.

However, I do not think that the moral problem is so easily solved. The main reason is that the latter alternative is not a plausible explanation of Smith's version of (2). Recall that Smith, very plausibly, argues that (2) should be restated to say that it is conceptually necessary that if an agent judges that it is right that she ϕ s, then she is motivated to ϕ *on condition that she is rational*. It seems reasonable to think that the fact that we classify an agent's judgment as a moral judgment only if she is accordingly motivated need not be explained by anything about the content of the judgment. It is merely a matter of how we choose to classify certain judgments. By contrast, it is much less plausible to maintain that the fact that we classify an agent's judgment as a moral judgment only if she is accordingly motivated *on condition that she is rational* has nothing to do with its content. It seems that the judgment needs to have a particular content which explains that an agent who makes it is guaranteed to be accordingly motivated only if she is rational. As we shall, see Smith provides such an account.

4. Smith's Solution to the Moral Problem

The solution Smith propose of the moral problem goes through three steps to be explained below. The general idea is this: Analyze rightness in terms of normative reasons. Analyze normative reasons in terms of what an agent would desire if she were rational. As part of this step, maintain that an agent is rational insofar as she has coherent desires and maintain that rational agents would 'converge' in their desires. Finally, analyze judgments about normative reasons as beliefs about what the agent would desire if she were rational in this sense. In this manner, it is thought that moral judgments can be shown to be both practical and objective at the same time as it is insisted that only desires can motivate.

The first step consists in a conceptually necessary claim about the connection between moral rightness and normative reasons:

(R) *Rationalism*: 'If it is [morally] right for agents to ϕ in circumstances C, then there is a [normative] reason for those agents to ϕ in C.'¹²

The second step amounts to a platitude—which in Smith's vocabulary amounts to a conceptually necessary claim—about normative reasons according to which they consist in facts about what an agent would desire under full rationality:

(P) *Platitude about Reasons*: 'What we have normative reason to do is what we would desire that we do if we were fully rational.'¹³

12 Smith (1994), p. 62.

13 Smith (1994), p. 150. Smith sometimes uses 'reason' in the singular without clarifying whether he has in mind a *pro tanto* reason or *pro toto* reason (all-things-considered strongest reason). Similarly, he sometimes uses 'desire' without clarifying whether he has

After having presented this connection between normative reasons and rational desiring, Smith offers an account of practical rationality. A fully rational agent (i) has no false beliefs; (ii) has all relevant true beliefs, and (iii) has the capacity to deliberate correctly.¹⁴ The last condition entails that the agent is able to imagine relevant facts and to desire in accordance with means-ends norms. Most importantly, an agent who has deliberated correctly has a set of desires that is 'systematically justified', which means that it is completely coherent and unified.

According to Smith, our concept of normative reasons entails that fully rational agents would converge in their desires about what to do. In conjunction with (P) this claim entails that all agents, given sameness in circumstances, have the same normative reasons, that is, normative reasons with the same content. Indeed, Smith maintains that if reasons were thought to be relative to actual desires, they would not be normative since they would be arbitrary, and error-theory about normative judgments would follow.¹⁵ In view of the fact that normative reasons have this particular nature, they can be said to be *categorical*: They are not relative to actual desires, and agents who find themselves in the same circumstances have the same normative reasons.

The last step in Smith's solution of the moral problem is to argue that there is a conceptually necessary implication from beliefs about normative reasons to rational desiring:

(I) *Implication from Normative Beliefs to Rational Desiring*: 'If an agent believes that she has a normative reason to ϕ , then she rationally should desire to ϕ '.¹⁶

in mind a desire or strongest desires. At the same time, it seems to be assumed that the strength of reasons corresponds to strength of desires in rational agents. It is plausible to think that these unclaritys and assumptions complicate his argumentation. For a discussion of Smith's view of the strength of reasons and desires, and of the relevance of the distinction between the rationally requiring and rationally permissive strength of reasons, see Gert (2008), pp. 1–23.

14 Smith's conception of full rationality is a modified version of Bernard Williams's account of this notion. See Williams (1981), pp. 101–105.

15 '[T]o suppose that our concept of what constitutes a rational justification could be radically relative in some way to the interests or desires of those who make claims about what is rationally justified, so that the considerations that rationally justify relative to that agent may fail to justify relative to another—is to suppose, quite incoherently, that something completely arbitrary—the mere fact that a particular agent who is making a claim about rational justification happens to have the contingent interests or desires that she happens to have—could in some way constitute a normative fact: a fact about rational justification. But this is incoherent because the only decisive point we can make about normativity is that arbitrariness, as such, always undermines normativity' (Smith (1997), p. 90).

16 Smith (1994), p. 148. As Geoffrey Sayre-McCord argues, the implication should be read: 'If an agent believes she has reason to ϕ then if she is rational she will desire to ϕ ' (Sayre-McCord (1997), p. 64).

In Smith's view, (P) entails (I) on the assumption that an agent believes that she has a normative reason to ϕ . He argues in the following way. If an agent believes that she has a normative reason to ϕ , it follows from (P) that she believes that she would desire to ϕ if she were fully rational. Accordingly, if the agent believes that she has a normative reason to ϕ , but does not desire to ϕ , she is irrational because she believes that she would desire that she ϕ s if she were fully rational, but nevertheless does not desire to ϕ . The agent's irrationality is said to consist in a special kind of incoherence between her attitudes:

Imagine two agents who believe that they would desire that they ϕ in C if they had a maximally informed, coherent and unified set of desires. One of these agents also desires that she ϕ s in C, but the other does not desire that he ϕ s in C. What can we say about the relative merits of these two agents' psychologies, straight away, given just what we've said? The obvious thing to say, it seems to me, is that the former psychology exhibits more in the way of coherence than the latter. *The latter agent fails to have a desire that he believes he would have if he had a maximally coherent set of desires, and this fact, all by itself, constitutes a kind of incoherence or disequilibrium in his overall psychological state.* The former agent's desires do not suffer any such disequilibrium or incoherence. Even if her belief is false, she still enjoys a sort of coherence, or equilibrium, simply in virtue of the fact that she has a matching desire to ϕ in C.¹⁷

We have thus far been concerned with normative reasons in general rather than moral reasons in particular.¹⁸ What makes normative reasons to *moral* reasons, what characterises this subset of normative reasons, is that they are demarcated by platitudes concerning the content of morality.¹⁹ These platitudes provide together with (R) and (P) Smith's analysis of moral rightness:

Moral Rightness: ' ϕ -ing in circumstances C is [morally] right if and only if we would desire that we ϕ in C, if we were fully rational, *where ϕ -ing in C is an act of the appropriate substantive kind.*'²⁰

We can now see how Smith aims to solve the moral problem.

First, (R) Rationalism and (P) the Platitude about Reasons are maintained to entail (I) the Objectivity of Moral Judgments. From (R) and (P), understood

17 Smith (1995b), p. 166. Emphasis added. Cf. Smith (199a5), pp. 126–127 and Smith (1994), pp. 177–179.

18 This has caused some misunderstanding of Smith's view; see Swanton (1996), pp. 155–160 and Smith (1996b), pp. 160–168.

19 Smith (1994), pp. 40–41, 183–184.

20 Smith (1994), p. 184.

in accordance with Smith's conception of full rationality, it follows that whether it is right that an agent ϕ s is determined by considerations that are not relative to her actual desires but determined by objective circumstances. That is, facts about what is morally right, a subclass of facts about normative reasons, are constituted by categorical requirements of rationality. Thus, (1) follows.

Second, (R) Rationalism and (P) the Platitude about Reasons are maintained to entail (2), the Practicality of Moral Judgments. We have already seen that Smith argues that (P) entails (I), the Implication from Normative Beliefs to Rational Desiring, on the assumption that an agent believes that she has a normative reason to ϕ . Now, from (R) and (P), together with this assumption, follows a moral analogue to (I): An agent who judges that it is morally right to ϕ —i.e. believes that she has a moral reason to ϕ —would be motivated to ϕ on condition that she is rational. Thus, (2) follows.

Lastly, the three steps—(R) Rationalism, (P) the Platitude about Reasons, and (I) the Implication from Normative Beliefs to Rational Desiring—are maintained to be compatible with (3) the Humean Theory of Motivation, the view that beliefs cannot motivate and that only desires are able to fulfil this function. Importantly, it is the claim that an agent's moral judgment consists in a belief about what she would desire if she were fully rational which explains that it is accompanied by a desire—motivation—to perform the action given that she is rational. Is not claimed that the moral belief in itself is motivating in insolation from any desire. Thus, the moral problem is solved.

In this paper, I will mainly be concerned with Smith's contention that moral judgments are both practical and objective. More precisely, I will argue that Smith's conception of (P), the Platitude about Reasons, is such that (2) the Practicality of Moral Judgments and (1) the Objectivity of Moral Reasons have not clearly been shown to follow. As a result, it might be doubted whether Smith has succeeded to solve the moral problem. Smith has developed his metaethical views after *The Moral Problem*.²¹ In what follows, I will primarily focus on Smith's view as it is formulated in the book, since the present issue of *Belgrade Philosophical Annual* is dedicated to the thirtieth anniversary of its publication.

5. Example Model and Advice Model

As was seen in the last section, (P) the Platitude about Reasons is a vital step in Smith's solution to the moral problem. We might start with considering a certain ambiguity in (P) identified by Smith in his seminal paper from 1995.²²

21 See in particular Smith (2009), pp. 98–125 and Smith (2012), pp. 309–331. In understanding Smith's view, I have learned a lot from David Brink's, David Copp's, and Geoffrey Sayre-McCord's comments on *The Moral Problem*: Brink (1997), pp. 4–32; Copp (1997), pp. 33–54, and Sayre-McCord (1997), pp. 55–83.

22 Smith (1995a), pp. 110–112, 125–129. For criticism of Smith's paper, see Johnson (1997), pp. 619–625.

The simplest reading of (P) is:

(P.1) *Example Model*: A (less than fully rational) agent has a normative reason to ϕ in a given situation if and only if she would desire to ϕ in that situation if she were fully rational.

According to (P.1), what an agent's fully rational self would desire to do forms an *example* to her less than fully rational self. However, this is not how (P) should be understood in Smith's view. What we have normative reasons to do, which Smith equates with what is *desirable* that we do, is not necessarily what we would desire to do if we were fully rational. As we actually are, we have beliefs and desires others than those we would have if we were fully rational. Accordingly, we might, as fully rational agents, be motivated to perform actions that we, as we actually are, have no reasons to perform. Conversely, we might, as we actually are, have reasons to perform actions that we would not, as fully rational agents, be motivated to perform. Smith has a well-known example that is designed to show why this is the case.²³

Squash Example: You have been defeated in a game of squash in a way that you find humiliating. If you were fully rational, you would have been calm and walked to your opponent to shake hands with him. However, as you are not fully rational, you are consumed by a desire to smash your opponent in the face with your racket.

In Smith's view, what it is desirable that you do is not to stride over and shake hands with your opponent, but to leave the court as soon as possible.

In Smith's view, (P) should instead be read as follows:

(P.2). *Advice Model*: A (less than fully rational) agent has a normative reason to ϕ in a given situation if and only if her fully rational self would desire that her less than fully rational self ϕ s in that situation.

According to (P.2), an agent's fully rational self *advises* her less than fully rational self about what to do. In the squash example, leaving the court is not what your fully rational self would desire to do, since she does not have a desire to smash your opponent in the face. However, walking away is what your fully rational self would desire that you, being less than fully rational, do, since it is the preferable alternative given your desire.

6. Doubt about the Advice Model

We might start with considering Smith's motivation for rejecting (P.1) the Example Model in favour of (P.2) the Advice Model. An intuitive objection against Smith's argument in relation to the squash example might be that it is

23 Smith (1995a), pp. 111–112.

desirable that you shake hands with your opponent and that this is the case quite independently of your desires.

One possible reply is that it is desirable that you do not walk to your opponent to shake hands with him but leave the court because it would have the most favourable consequences.²⁴ If you stride up to your opponent to shake hands, you will smash him in the face, but if you walk away that will not happen. However, it might be worried that the reply is not satisfactory. First, it seems unclear whether Smith is in the position to appeal to favourable consequences of actions without having a substantive view of rationality in place.²⁵ Second, it might be objected that what is desirable that you do is to shake hands with your opponent *without* smashing him in the face and that *this* would have the most favourable consequences. The sad fact that you will smash your opponent if you walk to him to shake hands is neither here nor there for what is desirable that you do or what has preferable results.

It might perhaps be protested that the objection betrays a misunderstanding of the example. One possible interpretation is the following. What is wrong with (P.1) is that it implies that what is desirable is that you do something that you could do if you were fully rational, but that you *cannot* do because you have desires you would not have if you were fully rational. That is, (P.1) conflicts with an analogue to 'ought implies can': It is desirable that an agent performs a certain action only if she could get herself to perform it.²⁶

In view of the extensive debate on 'ought implies can', it would be premature to assume any particular view about the dictum 'desirable implies can'. However, it is plausible to think that you in the relevant sense *can* walk to your opponent and shake hands with him without smashing him in the face because you would do so *if* you desire to and your desire to smash your opponent were not allowed to prevail. This suggestion gains support from the common idea that an agent could have performed a certain action insofar as she would have performed it if she had desired or willed to. Moreover, the fact that an agent has or lacks a certain desire is ordinarily not considered as an acceptable excuse for her performing an undesirable action or not performing a desirable action, in contrast to other excuses, such as physical obstacles or ignorance.

However, let us grant that Smith's argument against (P.1) and in favour of (P.2) applies in cases where (P.1) would entail that it is desirable that an agent performs an action she is unable to perform. Nevertheless, I think there is an argument against (P.2) that applies in other cases. We can think of cases where (P.2) entails that it is desirable that an agent performs a certain action,

24 Cf. Smith (1995a), p. 111.

25 I owe this point to Joshua Gert.

26 Cf. Johnson (1997), pp. 619–625.

but where it is desirable that she performs quite another action that she is able to perform. Consider a slightly modified version of the squash example:

Squash Example Modified: You are not such that you cannot avoid smashing your opponent in the face if you walk to him to shake hands. However, you have a strong desire to do so. In consideration of this desire, it is very probable that you will smash him in the face if you stride up to him to shake hands.

In view of your fully rational self's favourable epistemic position and her resulting information about probabilities, it seems reasonable to think that she would not desire that you stride up to your opponent to shake hands, but that you walk away. Moreover, in view of your fully rational self's epistemic position and information about probabilities, it seems that she would desire this even if, as things ultimately turn out, you walk to your opponent and shake hands with him without smashing him in the face. However, in this case it seems less plausible to maintain that what is desirable that you do is to walk away instead of striding up to your opponent to shake hands.

In the type of case under consideration, (P.2) does not get support from the 'desirable implies can' principle. We are considering cases with the following features: (i) An agent's fully rational self would desire that her less than fully rational self performs a certain action since the latter suffers from an imperfection as a result of lacking rationality; (ii) If her less than fully rational self did not suffer from this imperfection, her fully rational self would desire that her less than fully rational self performs quite another action, and (iii) Her less than rational self is able to perform another action. In this type of cases, (P.2) seems to give the wrong verdict. The reason is that what an agent's fully rational self would desire that her less than fully rational self do might depend on her less than fully rational self's actual desires at the same time as the latter is able to perform another action which we consider desirable.

It might be replied that your fully rational agent not only would predict that you would smash your opponent in the face. As she has no false beliefs and all relevant true beliefs, she would actually *know* what you would do. However, if we think of your fully rational self's epistemic position in this manner, it does not seem evident that she would have any desire concerning your actions. First, it might be asked whether your fully rational self would have any relevant desire concerning your actions if she already knows what you will do.²⁷ Smith is not explicit about whether knowledge of the future is included in a fully rational agent's 'all relevant true beliefs'. However, since it implies that a fully rational agent's desires depend on her beliefs, there is

27 This understanding of a fully rational agent's epistemic position would also be a problem for (P.1) the Example Model. Would you have a relevant desire concerning your actions if you already know what you will do?

some reason to think that knowledge of the future is not included.²⁸ Second, on Smith's view of desires in terms of 'direction of fit', desires essentially seem to be something that are directed towards an unknown future. According to this view, what characterizes a desire that *p* in contrast to a belief that *p* is their having different relations to the perception that not *p*. A belief that *p* is such that it would 'tend to go out of existence in the presence of a perception with the content that not *p*, whereas a desire that *p* tends to endure, disposing the subject in that state to bring it about that *p*.²⁹ If an agent knows that the world fits her desire that *p* or that it will fit her desire that *p*, it is plausible to hypothesize that her desire that *p* will vanish, corresponding to a belief that *p* which will go out of existence in the presence of a perception that not *p*.

It might be replied that an agent who desires that *p* and who comes to know that *p* will be the case still will continue to desire that *p* because she is aware that *p* will not realize unless she desires that *p*. That is, she is aware that her having a desire that *p* is causally necessary for *p* to realize. However, this is not the situation of an agent's fully rational self. An agent's fully rational self has a desire *about* what her less than fully rational self is to do. Thus, it is not an agent's fully rational self having a desire that is causally necessary for something to realize; it is her less than fully rational self having a desire which has this function.

7. The Example Model and the Implication from Normative Beliefs to Rational Desiring

We saw above that Smith argues as follows in order to establish (2) the Practicality of Moral Judgments: (P) the Platitude about Reasons entail (I) the Implication from Normative Beliefs to Rational Desiring on the assumption that an agent believes that she has a normative reason to perform an action. From (R) Rationalism and (P) together with this assumption follows a moral analogue to (I): An agent who judges that an action is right for her to perform—i.e. believes that she has a moral reason to perform it—is motivated—desires—to perform the action on condition that she is rational. Thus, (2) has been established. This line of argument means that insofar as it turns out that (P) entails implausible versions of (I), there is reason to doubt that (2) has been established.

In the last section, it was found that there are reasons to doubt whether Smith has shown that (P.2) the Advice Model is to be preferred to (P.1) the Example Model. We might then return to (P.1). However, I think it might be argued that (P.1) entails implausible versions of (I).

If we insert the account (P.1) provides of normative reasons in (I), this implication should be read in the following way:

28 See Smith (1994), pp. 156–157 and Smith (1995a), pp. 112–113.

29 Smith (1994), p. 115.

(I.1) If an agent believes that she would desire to ϕ if she were fully rational, then she would desire to ϕ if she were rational.

In considering (I.1), it can be seen that 'rational' in the consequent cannot plausibly be understood in exactly the same way as 'fully rational' in the antecedent. In that case, the consequent would stipulate a condition of the truth of the implication that might be false according to the antecedent, namely that the agent is fully rational. We should then consider the possibility that 'rational' in the consequent concerns a relevant part of being 'fully rational'.

According to Smith's conception of rationality, it follows that an important aspect of an agent being fully rational is that she has a completely coherent set of desires. In the long quotation from Smith above, we saw that his argument that (P) entails (I) appeals to a particular type of incoherence between attitudes.³⁰ In line with Smith's argument, it might then be maintained that an agent is rational only insofar as she desires to do what she believes that she would desire to do if she had a completely coherent set of desires. We get the following version of (I):

(I.2): If an agent believes that she would desire to ϕ if her set of desires were completely coherent, then she would desire to ϕ if she desires to do what she believes she would desire to do if her set of desires were completely coherent.

However, explicated in this way it is evident that the conception of rationality referred to in the consequent (the second string of words underlined) is not part of the conception of rationality referred to in the antecedent (the first string of words underlined). The kind of rationality referred to in the antecedent concerns a desire an agent has given that it is coherent with her other *desires*. The kind of rationality referred to in the consequent concerns a desire an agent has given that it is coherent with a *belief* about what desire she would have if her set of desires were coherent.

The fact that the antecedent and the consequent refer to different conceptions of rationality gives rise to two difficulties. First, as far as I see, we have not been offered a clear account of the contention that coherence between beliefs and desires is part of rationality or, for that matter, that beliefs can cohere with desires. The latter is perhaps also something that might be questioned on the view that desires and beliefs have different 'directions of fit'.³¹ Second, it seems that (I.2) involves a certain type of inconsistency. The conception of rationality referred to in the antecedent states that an agent is fully rational only if she has the desires she would have if her set of desires were

30 Smith (1995b), p. 166.

31 Sayre-McCord has raised a similar question. As far as I understand, Smith's response is more concerned with the coherence between desires than the coherence between beliefs and desires. See Sayre-McCord (1997), pp. 74–76 and Smith (1997), pp. 92–99.

completely coherent. However, the conception of rationality referred to in the consequent is in potential conflict with this notion. It states that an agent is rational only insofar as she has the desires that she believes she would have if she had a coherent set of desires. As a result, (I.2) involves two conceptions of rationality that make potentially conflicting demands: ‘Desire as you would if you had a completely coherent set of desires’ and ‘Desire as you believe you would if you had a completely coherent set of desires’³² The two demands will conflict with one another in cases where an agent has false beliefs about what she would desire if she had a completely coherent set of desires. Assume that an agent falsely believes that she would have a certain desire if she had a completely coherent set of desires. Assume further that she as a consequence of this belief comes to acquire the desire. The agent now has a desire which she rationally should have in accordance with the conception of rationality referred to in the consequent and she has thus become more rational on this conception. However, according to the conception of rationality referred to in the antecedent, it might be that she has become *less* rational because now she has a desire that she would not have if her set of desires were completely coherent.

8. The Advice Model and the Implication from Normative Beliefs to Rational Desiring

As we have seen, Smith’s main motivation for preferring (P.2) the Advice Model to (P.1) the Example Model seems to be that the former squares better with our intuitive conception of what is desirable. However, it was found that there are reasons to doubt this assumption. In the last section, it was argued that (P.1) entails implausible versions of (I) the Implication from Normative Beliefs to Rational Desiring. It might then be suspected that the motivation for preferring (P.2) to (P.1) rather should be found in the conviction behind (I). We might therefore return to (P.2) and consider its relation to (I). However, as we shall see, it might be argued that also (P.2) entails an implausible version of (I).

In a central passage in the paper from 1995, Smith argues that the relation between beliefs about normative reasons and rational desiring supports (P.2) in favour of (P.1):

Suppose, for instance, that you believe your fully rational self would desire to ϕ in the circumstances she faces; that this is the example she would set for you in her own word. Why should this have any effect at all on what you desire to do in the circumstances you face? [---] Coherence and unity do not argue in favour of acquiring a desire like hers because her example—marvelous though it is in the circumstances

32 Cf. Persson (1995), p. 151; Dancy (1996), p. 178, and Sayre-McCord (1997), pp. 55–83.

in which *she* finds *herself*—doesn't engage with the circumstances in which *you* find *yourself*. This is not the case if instead we interpret the requirement in terms of the Advice Model. For then what you have to *believe* is that *your fully rational self would want your less than fully rational self to ϕ in the circumstances your less than fully rational self actually faces*. Your fully rational self's advice engages with your predicament because it is precisely tailored to it. You may still say 'So what?', of course, but if you do you simply reveal that you are unable to accept good advice; you reveal the extent to which your psychology fails in terms of norms of coherence and unity that define a systematically justified psychology. You thus simply betray your own irrationality.³³

Thus, in effect the motivation for preferring (P.2) to (P.1) seems to come from an argument to the effect that (P.2) squares better with (I). In passing, it might be reflected that the contention seems troublesome in the context of Smith's overall argumentation. According to this line of argument, (P), together with the assumption that an agent believes that she has a normative reason, entails (I). However, if (I) is established because it follows from (P), it is questionable whether (I) can be used to support (P) or a particular version of it.

Nevertheless, I think it might be worried that (P.2) has essentially the same difficulty as (P.1): It entails an implausible version of (I). However, now things become a bit more complicated. According to (P.2.), it follows that an essential aspect of an agent being fully rational is that she, having a completely coherent set of desires, would desire that she, as she actually is, performs certain actions. In line with Smith's argument in the quotation above, it might then be maintained that an agent is rational only insofar as she desires to do what she believes that she, had she a completely coherent set of desires, would desire that she, as she actually is, do. The following convoluted version of (I) is the result:

(I.3) If an agent believes that she, had she a completely coherent set of desires, would desire that she, as she actually is, ϕ s, then she would desire to ϕ if she desires to do what she believes that she, had she a completely coherent set of desires, would desire that she, as she actually is, do.

However, it is doubtful whether (I.3) involves the relevant type of coherence. Consider the desire referred to in the antecedent (the first string of words underlined) and in the consequent (the second string of words underlined), respectively. In the first case, there is a desire which is the object of an agent's belief about what she would desire (if she had a completely coherent set of desires) that she do (as she actually is). That is, the desire which is the object

33 Smith (1995a), pp. 128–129. Emphasis added.

of the belief is an agent's hypothetical desire that she performs an action in a world where she is different from how she is in the world in which she has that hypothetical desire. In the second case, there is an agent's desire to perform an action as she is when she has the desire, in the actual world. The first kind of desire is a desire with the content 'that I do action x as I would be in that world'. The second kind of desire is a desire with the content 'to do action x as I am in this world'. The difference is indicated by the phrases 'desire that' and 'desire to'. In view of the fact that the contents of these desires are structured in different ways, it might be doubted that it is a matter of coherence of the requisite kind.

It should also be noticed that (I.3) is subject to the same type of problems as (I.2) noticed above. First, (I.3) refers to a conception of rationality that rests on the contention that desires can cohere with beliefs. Second, (I.3) involves two potentially conflicting demands of rationality: 'Desire as you, had you a completely coherent set of desires, would desire that you, as you actually are, do' and 'Desire as you believe that you, had you a completely coherent set of desires, would desire that you, as you actually are, do'. The two demands will conflict with one another in cases where an agent has false beliefs about what she, had she a completely coherent set of desires, would desire that she, as she actually is, do.

9. The Implication from Normative Beliefs to Rational Desiring and the Conception of Rationality

In the two proceeding sections, I have tried to argue that both versions of (P) the Platitude about Reasons—(P.1) the Example Model and (P.2) the Advice Model—entail implausible versions of (I) the Implication from Normative Beliefs to Rational Desiring. In view of the fact that (2) the Practicality of Moral Judgments is supposed to constitute a moral analogue to (I), there is reason to doubt that (2) has been established.

However, there might be a more direct manner of casting doubt on whether Smith has established (2). The reason is that (2), as originally conceived, and (I) seem to refer to different conceptions of rationality.

According to (2), if an agent judges that an action is right for her to perform then she is motivated to perform it on condition that she is rational. According to Smith's original conception of rationality, an agent is fully rational insofar as she has no false beliefs, all relevant true beliefs, is able to imagine relevant facts, and has a set of desires that is completely coherent and unified. The original conception of rationality is modelled to explain that an agent who judges that an action is right for her to perform might not be motivated to perform it if she suffers from common forms of irrationality, such as weakness of will or depression. For example, a depressed agent

might judge that an action is right for her to perform but nevertheless not be motivated to perform it in case she lacks some relevant true beliefs, cannot correctly imagine alternative actions available to her, or some of her desires are incoherent.

In the discussion above, we have found that the different interpretations of (P) entail different versions of (I). However, what is important in the present context is that the conception of rationality referred to in these versions of (I) differs from Smith's original conception of rationality as referred to in (2). Thus, in (I.3) the consequent entails that an agent is rational only insofar as she desires to do what she believes that she, had she a completely coherent set of desires, would desire that she, as she actually is, do.³⁴

The fact that (2) and (I) refer to different conceptions of rationality has implications for whether Smith has succeeded to establish (2).

First, an agent fulfilling the modified conception of rationality is compatible with her not desiring to do what she judges is right for her to do. According to (2), if an agent judges that an action is right for her to perform—i.e. believes she has moral reason to perform it—then she is motivated—desires—to perform the action on condition that she is rational. In Smith's view, an agent has a normative reason to perform an action if she would desire that she performs it if she were fully rational. As we have seen, according to the original conception of rationality an agent is fully rational insofar as she has no false beliefs, all relevant true beliefs, is able to imagine relevant facts, and has a coherent and unified set of desires. By contrast, according to the modified conception of rationality referred to in the consequent of (I.3) an agent is rational only insofar as she desires to do what she believes that she, having a fully coherent set of desires, would desire that she, as she actually is, do. An agent being rational on the modified conception is evidently compatible with her not desiring to do what she believes that her fully rational self would desire that she do. As a consequence, it is difficult to see that it has been established that if an agent judges that an action is right for her to perform, then she is motivated to perform it on condition that she is rational, according to Smith's modified conception of rationality.

Second, an agent fulfilling the modified conception of rationality is compatible with her suffering from common forms of irrationality, such as weakness of will and depression. Assume that an agent desires to do what she believes that she, having a coherent set of desires, would desire that she, as she actually is, do. However, this is compatible with the agent, for example, lacking relevant true beliefs, being incapable of imagining available actions correctly, and having incoherent desires. The modified conception of rationality does thus not rule out depression and other instances of practical irrationality that constitute the target of Smith's original conception of this

34 In the argument, I consider (I.3). However, similar results emerge if we instead consider (I.2).

notion. As a consequence, the modified conception of rationality fails to explain that an agent who judges that an action is right for her to perform might not be motivated to perform it if she suffers from some common form of irrationality.

Finally, and more generally, the two above considerations underline that the modified conception of rationality referred to in (I) differs from the original conception of rationality as referred to in (2). As a consequence, it is difficult to see that (2) constitutes a moral analogue to (I) in the manner Smith's argument for (2) seeks to establish.

10. The Advice Model and the Objectivity of Moral Judgments

As a final point, I think it can be argued that the fact that Smith decides for (P.2) the Advice Model over (P.1) the Example Model might have significant implication for whether (1) the Objectivity of Moral Judgments has been established.

We have seen that Smith argues as follows in order to establish (1): (R) Rationalism and (P) the Platitude about Reasons, understood in accordance with the original conception of rationality, entail that whether it is right for an agent to perform an action is determined by considerations that are not relative to her actual desires but determined by objective circumstances. Thus, moral judgments are categorical which means that (1) has been established.

However, in case (P) is understood in line with (P.2) it can be doubted that Smith's analysis of moral judgments entails that moral judgments are categorical.³⁵ According to (P.2), what constitutes a normative reason for a less than fully rational agent depends on what her fully rational self would desire that she do, given how she actually is. As a result, it seems that there is no guarantee that all less than fully rational agents have the same normative reason if they find themselves in the same circumstances. This is so since what their fully rational selves would desire that their less than fully rational selves do depends on the desires of the latter. Notice that it may still be granted that all fully rational agents would end up with having converging desires so that they would all desire that a less than fully rational agent performs a particular action in a given situation. However, there is no guarantee that they would desire that two less than fully rational agents who find themselves in the same circumstances perform the same action if the desires of these agents differ in any significant respect. It follows that what normative reasons agents have might be relative to their actual desires. And from this it seems to follow that moral judgments are *not* categorical: different actions might be right for different agents to perform, although they find themselves in the same circumstances. The consequence seems to be that (1) the Objectivity of Moral Judgments cannot be sustained: moral judgments are not objective.

35 For related worries, see Hubin (1999), pp. 355–361 and Sobel (1999), pp. 137–147.

It should be stressed, however, that in Smith's view what would follow is not that moral judgments are not categorical. Rather, what would follow is that all moral judgments to the effect that actions are right are false. As noticed above, Smith maintains that if reasons were thought to be relative to actual desires, they would not be truly normative as they would be arbitrary. As a consequence, error-theory about normative judgments would be the result.

An obvious reply to this line of argument is to maintain that two less than fully rational agents have reasons to perform the same action in the same circumstances if 'circumstances' include actual desires. However, it is difficult to see that this understanding of 'circumstances' would make the account less conditional and save the categoricity of moral judgments in any substantive sense. Moreover, it seems a bit strained to maintain that an agent's desires are part of the circumstances of her situation.

It might further be replied that my argument merely shows that the demand that an agent's moral reasons should not be relative to her actual desires is too strong and has to be modified. Consider the squash example. It might be argued that what it is desirable that you do depends on your actual desire to smash your opponent in the face. To demand that what is desirable that you do should not be relative to your actual desire would be to demand that you should do something that is not desirable, namely to shake hands with your opponent with the result that you smash him in the face. Alternatively, it might be argued that to demand that what is desirable that you do should not be relative to your actual desire would be to demand that you should do something that you cannot do, namely to shake hands and leave the court without smashing your opponent in the face.

These arguments have already been considered above and I will not repeat my discussion of them. However, it should be recalled that (P.2) entails that in some cases where an agent's fully rational self forms a desire about what her less than fully rational self should do, she is sensitive to her less than fully rational self's desires even if the latter is able to perform an action that is considered desirable. To the extent that (P.2) has this implication, it seems to follow that what is right for an agent to do might be relative to her actual desires.

11. Conclusion

In his seminal book *The Moral Problem*, Smith sets out to solve the moral problem. In essence, the problem is that three intuitively plausible claims seem incompatible: (1) the Objectivity of Moral Judgments, (2) the Practicality of Moral Judgments, and (3) the Humean Theory of Motivation. Smith argues that these claims can be reconciled when properly understood. In this paper, I have primarily been concerned with Smith's efforts to establish (1) and (2) by considering his conception of practical rationality understood as coherence

between attitudes. Smith argues that (1) and (2) follow from (R) Rationalism and (P) the Platitude about Reasons. However, it was found that there are reasons to doubt this contention. In considering Smith's understanding of (P) it was argued that it entails implausible versions of (I) the Implication from Normative Beliefs to Rational Desiring, of which (2) is assumed to be a moral analogue. As a result, there are reasons to doubt that (2) has been shown to follow. Moreover, it was argued that Smith's understanding of (P) is incompatible with moral judgments being categorical. As a result, there are reasons to doubt that (1) has been shown to follow. Considering that there are reasons to question whether Smith's has established that (1) and (2) follow from his analysis of the relevant claims, there are also reasons to question that he has demonstrated that they are compatible. Thus, it can be doubted that the moral problem has been solved.³⁶

References

- Arruda, C.T. (2019). 'What the Humean Theory of Motivation Gets Wrong.' *Journal of Philosophical Research* 44: 157–178.
- Björklund, F. et al. (2012). 'Recent Work on Motivational Internalism.' *Analysis* 72(1): 123–137.
- Brink, D. (1997). 'Moral Motivation.' *Ethics* 108(1): 4–32.
- Broome, J. (2013). *Rationality through Reasoning*. Oxford: Oxford University Press.
- Copp, D. (1997). 'Belief, Reason, and Motivation: Michael Smith's The Moral Problem.' *Ethics* 108(1): 33–54.
- Dancy, J. (1996). 'Real Values in a Humean Context.' *Ratio* 9(2): 171–183.
- Francén, R. (2020). 'Reconsidering the Meta-ethical Implications of Motivational Internalism and Externalism.' *Theoria* 86(3): 359–388.
- Gert, J. (2004). *Brute Rationality*. Cambridge: Cambridge University Press.
- Gert, J. (2008). 'Michael Smith and the Rationality of Immoral Action.' *Journal of Ethics* 12(1): 1–23.
- Hubin, D. (1999). 'Converging on Values.' *Analysis* 59(4): 355–361.
- Johnson, R. N. (1997). 'Reasons and Advice for the Practically Rational.' *Philosophy and Phenomenological Research* 57(3): 619–625.
- Kiesewetter, B. (2017). *The Normativity of Rationality*. Oxford: Oxford University Press.
- Kolodny, N. (2005). 'Why be Rational?.' *Mind* 114(455): 509–563.

36 I am grateful to Joshua Gert for helpful comments on an earlier version of this paper.

- Korsgaard, C. (1996). 'Skepticism about Practical Reason.' In *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, pp. 188–221.
- Lillehammer, H. (1999). 'Analytical Dispositionalism and Practical Reason.' *Ethical Theory and Moral Practice* 2(2): 117–133.
- Lord, E. (2018). *The Importance of Being Rational*. Oxford: Oxford University Press.
- Markovits, J. (2014). *Moral Reasons*. Oxford: Oxford University Press.
- Parfit, D. (2011). *On What Matters*. Oxford: Oxford University Press.
- Persson, I. (1995). 'Critical Notice of Michael Smith: The Moral Problem.' *Theoria* 61(2): 143–158.
- Sayre-McCord, G. (1997). 'The Metaethical Problem.' *Ethics* 108(1): 55–83.
- Schroeder, M. (2007). *Slaves of the Passions*. Oxford: Oxford University Press.
- Smith, M. (1989). 'Dispositional Theories of Value.' *Proceedings of the Aristotelian Society*, Supplementary Volume 63(1): 89–174.
- Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.
- Smith, M. (1995a). 'Internal Reasons.' *Philosophy and Phenomenological Research* 55(1): 109–131.
- Smith, M. (1995b). 'Reply to Ingmar Persson's Critical Notice of The Moral Problem.' *Theoria* 61(2): 159–181.
- Smith, M. (1996a). 'Internalism's Wheel.' In *Truth in Ethics*, ed. Brad Hooker. Oxford: Blackwell, pp. 277–302.
- Smith, M. (1996b). 'Normative Reasons and Full Rationality: Reply to Swanton.' *Analysis* 53(3): 160–168.
- Smith, M. (1997). 'In Defense of The Moral Problem: A Reply to Brink, Copp and Sayre–McCord.' *Ethics* 108 (1): 84–119.
- Smith, M. (1998). 'Ethics and the A Priori: A Modern Parable.' *Philosophical Studies* 92(1–2): 149–172.
- Smith, M. (2009). 'Desires, Values, Reasons, and the Dualism of Practical Reason.' *Ratio* 22(1): 98–125.
- Smith, M. (2012). 'Agent and Patients.' *Proceedings of the Aristotelian Society*, Supplementary Volume 112(3): 309–331.
- Sobel, D. (1999). 'Do the Desires of Rational Agents Converge?' *Analysis* 59(3): 137–147.
- Strandberg, C. (2007). 'Externalism and the Content of Moral Motivation.' *Philosophia* 35(2): 249–260.

- Strandberg, C. (2011). 'The Pragmatics of Moral Motivation'. *The Journal of Ethics* 15(4): 341–369.
- Strandberg, C. (2012). 'A Dual Aspect Account of Moral Language'. *Philosophy and Phenomenological Research* 84(1): pp. 87–122.
- Strandberg, C. (2013). 'An Internalist Dilemma—and an Externalist Solution'. *The Journal of Moral Philosophy* 10(1): 25–51.
- Strandberg, C. (2024). 'Neo-Humean Rationality and Two Types of Principles'. *Analytic Philosophy* 65(2): 256–273.
- Swanton, C. (1996). 'Is the Moral Problem Solved?'. *Analysis* 53(3): 155–160.
- Tresan, J. (2006). 'De Dicto Internalist Cognitivism'. *Noûs* 40(1): 143–165.
- Williams, B. (1981). 'Internal and External Reasons'. In *Moral Luck*. Cambridge: Cambridge University Press, pp. 101–113.