

# Moral Reality

A Defence of Moral Realism

Caj Strandberg

Department of Philosophy  
Lund University  
2004

© Caj Strandberg  
ISBN 91-628-6257-X  
Printed by Media-Tryck, Lund University  
2004

# Contents

<b>Acknowledgements</b>	i
<b>Chapter 1. Realism in Meta-ethics</b>	
1. Introduction	1
2. Locating Moral Realism	2
3. Moral Realism in Terms of Irreducibility—A Defence	3
<i>The minimal conception</i>	
<i>Moral realism in terms of mind independence</i>	
<i>Moral realism in terms of literal construal</i>	
<i>Moral realism in terms of irreducibility</i>	
4. Moral Realism and Its Opponents	18
<i>Cognitivism contra non-cognitivism</i>	
<i>Success-theory contra error-theory</i>	
<i>Realism contra reductionism</i>	
<i>Realism and naturalism</i>	
5. Some Arguments against Non-cognitivism and Error-theory	30
6. Outline of the Argument	34
7. Concluding Remarks	37
<b>Chapter 2. Analytic Reductionism and the Open Question Argument</b>	
1. Introduction	39
2. Analytic Reductionism	39
3. An Amended Version of the Open Question Argument	40
4. Defence of the First Part of the Argument	46
5. Defence of the Second Part of the Argument	54
6. Concluding Remarks	59
<b>Chapter 3. Synthetic Reductionism and the Two Questions</b>	
1. Introduction	60
2. Natural Kind Terms	61
3. Synthetic Reductionism	64
4. The First Question	71
5. The Second Question	77
6. Concluding Remarks	80

## **Chapter 4. Realism and the Realist Formula**

1. Introduction	82
2. The Realist Formula	83
3. Reusing Synthetic Reductionism	87
4. Concluding Remarks	90

## **Chapter 5. Explaining Moral Disagreement**

1. Introduction	91
2. The Argument from Moral Disagreement	92
3. Types of Moral Disagreement	94
4. Moral Disagreement and Error	95
5. Widespread Moral Disagreement	98
6. Persistent Moral Disagreement and the Phenomenon of Normative Divergence	101
7. The First Reductionist Explanation of Persistent Moral Disagreement	105
<i>Analytic reductionism</i>	
<i>Synthetic reductionism</i>	
8. The Second Reductionist Explanation of Persistent Moral Disagreement	114
9. A Note on Relativism	117
10. A Realist Account of Normative Divergence	121
11. A Realist Explanation of Persistent Moral Disagreement	124
12. Explaining the Persistence of Persistent Moral Disagreement	131
13. Concluding Remarks	132
<i>Appendix. Persistent Moral Disagreement and Unknowable Moral Facts</i>	133

## **Chapter 6. Explaining Moral Reason**

1. Introduction	136
2. The Moral Reason Principle	138
3. Reductionism on the Moral Reason Principle	141
4. Realism on the Moral Reason Principle	143
5. 'Why Should I Do What Is Right?'	144
6. Three Objections	155
7. Analogous Normative Questions	158
8. Other Normative Questions	164
9. A Note on Rationalism and Realism	165
10. Concluding Remarks	168

<b>Chapter 7. Explaining Moral Motivation</b>	
1. Introduction	169
2. Internalism and Externalism	170
3. The Internalist Argument	173
4. Internalist Considerations	177
5. The Amoralist and Her Critics	183
6. Advanced Amoralists	186
7. Externalist Explanations of Internalist Considerations	192
8. Smith's Fetishist Argument	201
9. Reductionism, Realism and Moral Motivation	212
<i>Analytic reductionism</i>	
<i>Synthetic reductionism</i>	
<i>Realism</i>	
10. Concluding Remarks	218
<b>Chapter 8. Moral Dependence</b>	
1. Introduction	220
2. The Importance of Moral Dependence	221
3. General Dependence	226
4. Realist Dependence	229
5. The Realist Formula and Supervenience	239
6. Strong or Weak Supervenience?	240
7. Blackburn: Strong or Weak Supervenience—a Dilemma	242
8. Kim: Strong or Weak Supervenience—another Dilemma	249
9. Mackie's Arguments from Queerness	254
10. Concluding Remarks	265
<b>Chapter 9. Moral Explanations</b>	
1. Introduction	266
2. An Outline of Harman's Argument	267
3. The Causal Criterion	268
4. The Best Explanation Criterion	272
5. A Dilemma for Harman's Argument	281
6. The Pragmatic Criterion	284
7. Three Conceptions of the Causal Powers of Moral Properties	288
8. Concluding Remarks	296
<b>References</b>	299



## Acknowledgements

I have received a tremendous amount of help from a large number of people in writing this thesis. At one point, I toyed with the idea of inserting footnotes in the text where I would express my gratitude for each comment I have benefited from. But it turned out to be impossible—there would simply be too many of them. Although this put some doubt in my mind about my own contribution to the text, it made me also realise that philosophy may not be such a lonely enterprise after all. So I would like to take the opportunity to thank my ‘co-writers’ here instead.

The person to whom I owe most gratitude is my supervisor Ingmar Persson. Ingmar was the one who suggested to me that moral realism is best understood as the view that moral properties are irreducible, an idea that finally got me start writing and which has provided the shape to this thesis. As a supervisor, Ingmar has been very considerate and generous with his time and wide-ranging philosophical knowledge. His constructive and thought-providing comments during the writing process helped me improve the text significantly. He has also—which perhaps is most admirable—been surprisingly patient with my slow-wittedness and frequent confusion. Another person to whom I am very grateful is Wlodek Rabinowicz. Wlodek’s passionate interest in philosophy and his astonishing insight into numerous philosophical fields is a source of inspiration to me and other people at the philosophy department in Lund. He has generously provided suggestive and detailed comments on my drafts, which helped to greatly improve the text. I have also benefited a great deal from extensive written commentaries on individual chapters by Åsa Andersson, Sven Nyholm and Folke Tersman. The participants of the Higher Seminar in Practical Philosophy in Lund have provided me with innumerable helpful comments on the various drafts I have presented there; I am particularly grateful to David Alm, Åsa Andersson, David Bengtsson, Johan Brännmark, Dan Egonsson, Jeanette Emt, Ylva von Gerber, Lena Halldenius, Victoria Höög, Magnus Jiborn, Mats Johansson, Andreas Lind, Sven Nyholm, Jonas Olson, Erik Persson, Björn Petersson, Toni Rønnow-Rasmussen and Daniel Svensson for useful suggestions. Parts of the text have been presented at

Filosofidagarna ('Philosophy Days'—national conference on philosophy) in Gothenburg, Stockholm and Linköping, and at the Higher Seminar in Practical Philosophy at Uppsala University. On these occasions, I have particularly benefited from valuable comments by Gunnar Björnsson, Jonas Olsson, Jan Österberg and Folke Tersman. During a stay as Recognised Student at the University of Oxford, I had interesting discussions about the open question argument with James Griffin. I am also grateful to Jonathan Dancy and Michael Smith for inspiring discussions about moral motivation and other meta-ethical issues.

But life is not merely philosophy—not even at a philosophy department. Together with Jonas Josefsson I have had the privilege of organising two conferences in philosophy. Although the conferences demanded more work than I first realised, they also provided a welcome contrast to tricky meta-ethical problems. I would like to thank Jonas, Wlodek, Lena Gunnarsson, Johannes Persson and Ann Tobin for making the work with the conferences so pleasant.

I have received generous financial support from Anders Karitz' stiftelse, Birgit och Gad Rausings stiftelse för humanistisk forskning, Crafoordska stiftelsen, Dagmar Perssons fond, Humanistiska fakulteten vid Lunds universitet, Knut och Alice Wallenbergs stiftelse, Kungliga vetenskapssamfundet i Lund, Stiftelsen Erik och Gurli Hultengrens fond för filosofi, Stiftelsen fil. dr. Uno Otterstedts fond, STINT (The Swedish Foundation for International Cooperation in Research and Higher Education). I am deeply grateful to all these institutions for making my work possible. I would also like to thank Dennis Brice who undertook to check my English in the final manuscript at short notice.

Last, but not least, I would like to thank all my wonderful friends. Although I have tried not to trouble them with my philosophical problems, they have meant more to my work than they can imagine. I am especially grateful to Jonas Rendahl for having solved the difficulties with my computer so readily. I dedicate this book to my parents, Annabella and Leif Strandberg, who have always encouraged and supported me.

C. S.



# Chapter 1

## Realism in Meta-ethics

### 1. Introduction

One significant human activity is to pass estimations of various sorts. We appraise actions, persons and other objects from a number of perspectives, such as law, morality, prudence and rationality. The estimations that play the most important role for our relations to other people are presumably the moral ones. We say things like ‘It’s right to help people in need’, ‘I think it’s wrong of Amanda to betray her husband’, ‘She’s a typically good person’ or ‘She showed how bad she is when she said all those terrible things about him’. Admittedly, many moral sentences are not as plain as these, but they can be taken to imply such sentences.

We may ask a number of philosophically intriguing questions in relation to moral sentences. To try to answer such questions is one of the enterprises of meta-ethics. Many moral sentences appear to be used to ascribe moral properties to objects. We may start to ask whether they, like other sentences of a similar form, are liable to truth and falsehood:

Are moral sentences capable of being true and false?

If we answer this question affirmatively, we may continue by asking whether there exist any moral rightness, wrongness, goodness and badness so that some such sentences are true:

Are there any moral properties that make certain moral sentences true?

If we also answer the second question affirmatively, it becomes relevant to ask about the nature of moral properties. There are various questions of this kind. However, one of the most essential is whether moral properties are different from other kinds of properties so that there is a distinct moral aspect of reality. We might formulate this question thus:

Do moral properties constitute a separate kind of properties?

To answer this question affirmatively is to deny that moral properties can be reduced to non-moral properties.

The meta-ethical view that answers these three questions affirmatively I will call *moral realism*. The main purpose of this thesis is to defend moral

realism thus understood. One line of argument will be that if we accept moral realism, we are better suited to account for certain important meta-ethical issues. As a part of this argument, I will maintain that answering the third question affirmatively makes it possible to answer the two first questions affirmatively. That is, I will argue that if we conceive of moral properties as making up a separate kind of properties, and hence as irreducible, we are able to sustain the view that moral sentences have truth-value and that there are moral properties that make some of them true. Moreover, I will defend moral realism against certain objections. However, I will also indicate significant issues where this defence is insufficient and moral realism is in need of further examination. This means that my defence of moral realism is limited in important respects.

In this chapter, I set the ground for the subsequent investigation. Following some preparatory remarks in the next section, I defend in section **3** a conception of moral realism, according to which this view should be understood in terms of moral properties being irreducible. In section **4**, I characterise the various claims that are involved in moral realism and their respective opposites. In section **5**, I consider very briefly some arguments against two of these views: non-cognitivism and error-theory. In section **6**, I give an outline of the argument of the remainder of the investigation.

## *2. Locating Moral Realism*

As I indicated in the last section, I take the defining feature of moral realism to be the claim that there are irreducible moral properties. ‘Realism’ is however a term that is used in numerous different senses. The senses in which it is used in ordinary language differ from the more or less technical ones it has acquired in various disciplines. The term has also acquired a number of senses within philosophy, not least in meta-ethics. Against this background it can justifiably be asked why we should understand moral realism in the way I advocate.

In response to this question, it may first be pointed out that to conceive of realism in terms of irreducibility is to understand it in accordance with *one* of the established senses of ‘realism’ in philosophy, meta-ethics included. Accordingly, it seems, at least *prima facie*, as legitimate to use the term in this sense as in some other of its established senses.

Another relevant consideration concerns the importance of moral realism. Irrespective of how meta-ethicists understand moral realism, they generally agree that the issue of whether this view is correct is important in meta-ethics. Conceiving of moral realism in terms of irreducibility preserves, I think, this conviction. As I indicated in the last section, one of the fundamental questions to ask about moral properties is whether they make up a distinct kind of properties. The importance of this question will be confirmed below.

Moreover, I think it can be argued that there are reasons to conceive of moral realism in the way I advocate rather than in some of the other ways current in the literature. In the next section, I will provide some arguments to this effect.

To facilitate this discussion, we may start by trying to localise where moral realism is to be found among the various philosophical claims that are called ‘realism’. It should first be noticed that the philosophical claims with this label have different ‘objects’ as their subject matter. There are thus realist claims about a number of notions which concern various ontological, semantic and epistemological issues. For example, there are realist claims about such grand notions as truth, properties in general, possible worlds and the external world. That there are realist claims about a wide range of notions helps to explain why ‘realism’ has acquired so many different senses. However, it is also striking that many of the realist claims have as their main subject matter certain kinds of entities that are discussed in particular fields of philosophy, such as aesthetics, the philosophy of mind and ethics. These claims concern ontological issues since they deal with the nature of the entities in question. Thus, there are realist claims that have as their main subject matter the nature of aesthetic properties, mental properties, moral properties, and so forth.<sup>1</sup> I will call this subclass of realist claims *ontological realist claims*. It is among ontological realists claims that we find moral realism.

### 3. *Moral Realism in Terms of Irreducibility—a Defence*

We may now continue with inquiring how moral realism should be understood.

---

<sup>1</sup> I say that the mentioned claims have these issues as their ‘main’ subject matter, since they may involve notions that are not ontological, e.g. various semantic notions. Moral realism is an example of such a view.

Although ‘realism’ is used in ordinary language, it seems in philosophy to have acquired more or less technical senses that differ from its ordinary senses. Therefore, in understanding these senses of the term, we have limited help from the knowledge we have in our capacity as competent language users, and consequently, cannot expect to reach a straightforward analysis of the term that captures this knowledge. This also holds for ‘moral realism’, and it seems therefore preferable to aim at finding a conception of moral realism that answers to some reasonable desiderata. While we have limited help from our knowledge as competent language users in developing a conception of moral realism, other considerations are more helpful. In the literature there are various conceptions of ontological realist claims and moral realism which characterise these views in such a way that their correctness is of philosophical significance. These considerations are relevant to the conception of moral realism that is to be preferred, as we will see.

There are, I think, three relevant desiderata for a successful conception of moral realism.

(i) A conception of moral realism should extend to alternative conceptions of moral realism.<sup>2</sup> Ideally, it should comprise these conceptions. It could, however, be difficult to find a conception which fulfils the desideratum in such a strong way. A conception may then fulfil it to a lesser degree, e.g. by explaining these conceptions. One rationale of this desideratum is that a conception of moral realism should make possible comparisons with the assertions of various meta-ethicists in relation to this view. Another rationale concerns the importance of moral realism. As mentioned above, whether moral realism is correct is generally thought to be an important issue. I also pointed out that various conceptions of moral realism characterise it in such a way that the correctness of this view is of philosophical significance. It is therefore desirable that the conception we choose extends to alternative conceptions.

(ii) A conception of moral realism should be an instance of a conception of ontological realist claims that extends to alternative conceptions of such claims.<sup>3</sup> This desideratum is an analogue to (i). As suggested above, and as we will see more clearly below, a certain

---

<sup>2</sup> Cf. Sayre-McCord (1988 (1986)), p. 5.

<sup>3</sup> Cf. Sayre-McCord (1988 (1986)), p. 5.

conception of moral realism is generally an instance of a certain conception of ontological realist claims. To fulfil this desideratum in an ideal way, a conception of moral realism should be an instance of a conception of ontological realist claims, which in turn should comprise alternative conceptions of such claims. However, as it might be difficult to find a conception of moral realism that fulfils this desideratum in such a strong way, it might only be possible to fulfil to a lesser extent. The rationales of this desideratum are counterparts to the rationales of desideratum (i). A conception of moral realism should make possible comparisons with ontological realist claims made in other fields of philosophy. Moreover, conceptions of ontological realist claims concern important philosophical issues, and it is therefore reasonable that they should have a counterpart in moral realism.

If a conception of moral realism fulfils desiderata (i) and (ii), it will provide an appealing coherence to debates surrounding moral realism. Ideally, a conception of moral realism comprises alternative conceptions of this view, and this conception of moral realism is in turn an instance of a conception of ontological realist claims that comprises alternative conceptions of such claims.

(iii) A conception of moral realism should be useful in discussions regarding meta-ethical issues, and, in particular, it should promote clarity when such issues are discussed.

As far as I understand, there are four main ways of conceiving moral realism. In the following, I will briefly indicate that the first three conceptions fail to satisfy the desiderata and then argue that the last conception fulfils the desiderata to a satisfactory extent. However, to provide a fully satisfactory defence of a conception of moral realism would require discussions of several intricate philosophical problems. This is unfortunately something I cannot do here.

### *The minimal conception*

According to the minimal conception of moral realism, this view comprises two conditions: (1) moral sentences have truth-value, and (2) there exist moral properties that make some moral sentences true.<sup>4</sup> As far as I understand, it is uncontroversial that these two conditions form part of

---

<sup>4</sup> See e.g. Blackburn (1993 (1973)), pp. 111–112, and Smith (1991), p. 402.

moral realism. The question is rather whether moral realism comprises some additional condition or conditions.

One difficulty with this conception is that it fails with regard to desideratum (i); that is, it fails to extend to alternative conceptions of moral realism. This point can be illustrated by an example of a meta-ethical position that fulfils the two conditions of the minimalist claim. Consider a view according to which moral sentences have truth-value, there is a moral property of actions, moral rightness, that make some of the sentences true, and this property consists in being desired by some human creature. It is highly doubtful whether we would say that this view qualifies as an instance of moral realism.<sup>5</sup> This example indicates that there are meta-ethical views which fulfil the two conditions contained in the minimal conception of moral realism but do not qualify as instances of moral realism on established conceptions of this view. It seems for example incompatible both with the conception according to which moral realism characterises moral properties as mind independent and the conception according to which moral realism construes moral sentences literally. (I will return to these conceptions below.)

Another difficulty with the minimal conception is that it fails with regard to desideratum (ii); that is, it is not part of a conception of ontological realist claims that extends to alternative conceptions of such claims. One example that illustrates this point can be brought from the philosophy of science: operationalism. On this view, scientific sentences have truth-value and there is a kind of ‘entities’—observable results of scientific operations—that make some of them true. However, operationalism is widely considered to be a non-realist position. This example indicates that there are views in other philosophical fields which fulfil counterparts to the two conditions of the minimal conceptions but which are not ontological realist claims according to alternative conceptions of such claims. It seems for example to run counter to the conception according to which ontological realist claims characterise the objects at issue as mind independent and to the conception according to which such claims characterise objects ‘at face value’. Given this difficulty, it is not surprising that the minimalist conception does not seem to have any counterparts as ontological realist claims.

---

<sup>5</sup> Cf. Railton (1996a), pp. 67–68, and Sayre-McCord (1988 (1986)), pp. 18, 22.

These difficulties show that the minimalist conception does not state a sufficient condition for moral realism. Hence, at least one more condition has to be added.

*Moral realism in terms of mind independence*

According to one influential conception of ontological realist claims, such claims characterise the kind of entities in question as mind independent.<sup>6</sup> How this exactly should be understood is not entirely clear. However, the central idea seems to be that an ontological realist claim characterises a certain kind of entities as objective in the sense that their nature is not conditioned by (the content of) our mental attitudes and other mental states. Thus, these entities are not the way they are just because we have certain observations, beliefs, desires or feelings, etc. Applied to moral realism, this means that the following condition should be added to (1) and (2): (3) moral properties are mind independent.<sup>7</sup> For example, the moral rightness of actions does not consist their being desired or their being objects of certain beliefs.

One difficulty with this conception is that it fails with respect to desideratum (ii). Admittedly, to conceive of moral realism in terms of mind independence might extend to conceptions of ontological realist claims as regards some fields of philosophy. The example from the philosophy of science above, operationalism, can perhaps serve as an illustration. But in relation to ontological realist claims as regards at least one field, the philosophy of mind, the conception under consideration does not fulfil this desideratum. As an illustration, consider Cartesian dualism. Very roughly, this might be understood as the view that there are two substances, the mental and the physical, which have separate existences. If we conceive of ontological realist claims in the indicated way, we would have to classify Cartesian dualism as a non-realist claim, since mental properties on this view are not mind independent; indeed, they consist, so to speak, in mere mind. But this seems mistaken because, if anything, Cartesian dualism is an

---

<sup>6</sup> The literature on realism as mind independence is vast, but for some instructive examples, see Boyd (1984), pp. 41–42; Devitt (1991), pp. 14–17, and Grayling (1987), pp. 25–27. Michael Dummett is known for understanding realist theories about truth in terms of mind independence; see e.g. Dummett (1982), pp. 55–112.

<sup>7</sup> For accounts of moral realism (partly) in terms of mind independence, see e.g. Dancy (1986), pp. 167–168; Miller (2003), p. 4; Pettit (2002), p. 52; Quinn (1978), pp. 257–258; Sturgeon (1986b), pp. 44–45, and Timmons (1999), p. 16.

example of realism in the philosophy of mind. However, this is only an illustration of a general difficulty with this conception of ontological realist claims. The general problem is that it cannot represent *any* view that has as its subject matter something related to mind as an ontological realist claim. Like others, Elliott Sober argues that realism should not be understood in terms of mind independence since the entities that are the subject matter of a certain view might be mental properties, in which case it would be incorrect to claim that they are independent of the mental.<sup>8</sup>

This point can be generalised even further. Suppose it is claimed that a view of a certain kind of entities qualifies as an ontological realist claim only if it characterises these entities as independent of X. Consider now a view that concerns X. No such view could be classified as an ontological realist claim as it can hardly be claimed to characterise X as independent of X. This suggests that it is mistaken to understand ontological realist claims in terms of independence of any *particular* phenomenon, mental or other.

It might be objected that this argument rests on a misunderstanding of how this conception of ontological realist claims comprehends the notion of mind independence. According to the notion of mental independence appealed to in this objection, such a claim states that the kind of entities in question is independent of mental states that concern this very kind of entities, particularly *beliefs about* this kind of entities. For example, an ontological realist claim about mental properties states that such properties are independent of beliefs about them. Accordingly, on this view moral realism states that moral properties are independent of beliefs about moral properties.<sup>9</sup> Understanding ontological realist claims in this way is not vulnerable to the objection mentioned above in relation to Cartesian dualism. On this conception, Cartesian dualism would be classified as a realist ontological claim, since this view characterises mental states as independent of what we believe about them. However, there are at least three difficulties with this view. First, it fails with regard to desideratum (i). The reason is that there are meta-ethical positions which claim that there

---

<sup>8</sup> Sober (1982a), pp. 370–371. That mind independence for this reason fails to provide a general criterion of the distinction between realist and non-realist claims is a recurrent argument in the literature; see e.g. Alston (1996), p. 73; Railton (1996a), pp. 51, 57, and Sayre-McCord (1988 (1986)), pp. 5–6, 15.

<sup>9</sup> Some of the authors who understand moral realism in terms of mind independence indicate, more or less explicitly, that they conceive of mind independence in this way; see e.g. Railton (1996a), p. 57.



exist moral properties that are independent of our beliefs about them but which do not seem to qualify as instances of moral realism. The view that moral rightness consists in being desired by some human creature can serve as an example. Second, for a corresponding reason it fails with regard to desideratum **(ii)**. The reason is that there are claims about a certain kind of entities which characterises the entities as independent of beliefs about them but hardly qualifies as ontological realist claims. Operationalism and behaviourism can perhaps serve as examples. Third, this view of mind independence seems to lead to a certain kind of regress. According to this view, a claim about a kind of entities X qualifies as an ontological realist claim in so far as it characterises X as independent of beliefs about X. Consider now a claim about *beliefs* about X. A claim about *these* beliefs qualifies as an ontological realist claim only if it characterises beliefs about X as independent of beliefs about beliefs about X, etc. To stop this regress, there has to be a certain kind of beliefs about which there is no ontological realist claim, in which case a difficulty similar to the one noted above emerges.

Let us return, then, to the original way of understanding moral realism in terms of mind independence. There is a further difficulty with this view that is worth mentioning: it fails with respect to desideratum **(iii)**; that is, it is not serviceable in discussions on meta-ethical issues. This difficulty can be illustrated by two examples of meta-ethical positions. Consider the view mentioned above according to which moral rightness of actions consists in their being desired by some human being. This view clearly does not characterise moral rightness as mind independent. Consider next a view which claims that moral rightness of actions consists in their being desired by a rational creature, where this is understood in a strict way, implying, for example, that the creature does not have any relevant false beliefs, all relevant true beliefs and a coherent set of desires.<sup>10</sup> Does this view characterise moral rightness as mind independent? In one sense it does, since the property on this view is independent of what people actually desire. In another sense it does not, since it makes use of mental concepts. This illustrates that it is difficult to make a clear distinction between views that characterise moral properties as mind independent and those that do not. Of course, we might define mind independence so exactly that each given view

---

<sup>10</sup> See Smith (1994), esp. chap. 5.

can be classified as characterising moral properties as mind independent or not. However, such a definition runs the risk of becoming arbitrary. For example, I think it would be arbitrary to say, without qualification, that the second position characterises moral rightness as mind independent; in one way it does, in another it does not. These considerations suggest, I think, that it is difficult to understand mind independence in a clear and non-arbitrary way that can be used to decide whether a certain view characterises a moral property as mind independent or not. As a result, on this conception it is difficult to draw a clear and non-arbitrary distinction between views that are instances of moral realism and those that are not. In addition, the examples above indicate that mind independence comes in degrees, that a certain meta-ethical view can be more or less mind independent. For these reasons, it does not seem helpful to use this notion to draw the line between moral realism and other meta-ethical views. In particular, it does not seem to promote clarity in discussions about meta-ethical issues.

It might be objected that on the conception that moral realism characterises moral properties as mind independent, moral realism should be understood to state that such properties are *totally* independent of mind. On this view, in order for moral properties to qualify as mind independent, these properties, and the non-moral properties in virtue of which objects have moral properties, should not be characterised in terms of any mental concepts whatsoever. If mind independence is understood in such a strict way, there is a clear and non-arbitrary understanding of mind independence, and the difficulty just mentioned is avoided. However, this understanding is presumably too strong, since quite few views of moral properties would classify as moral realism according to it. Consequently, it does not satisfy desideratum (iii). Furthermore, it fails to fulfil the other two desiderata.

#### *Moral realism in terms of literal construal*

According to another conception of ontological realist claims, such claims characterise the kind of entities in question ‘at face value’. What this exactly should be taken to mean is not entirely clear, but the basic idea seems to be that an ontological realist claim characterises the entities in question in a

way that conforms with our ordinary notion of them.<sup>11</sup> This means that there is a limit to how surprising the characterisation of the entities can be in order to qualify as realist, since, if it is very surprising, it presumably does not comply with our ordinary notion of these entities.<sup>12</sup> In the meta-ethical literature, a similar conception has been formulated as the view that moral realism states that some moral sentences are true when interpreted literally.<sup>13</sup> I take this to imply that these sentences are interpreted in such a way that the moral properties they ascribe to objects are understood ‘at face value’.<sup>14</sup> On this view, the following condition should thus be added to (1) and (2): (3’) some moral sentences are true when literally construed.

Conceiving of moral realism in this way seems to confront a difficulty in relation to desideratum (iii). According to this conception, in order for a meta-ethical view to qualify moral realism, it has to construe some moral sentences literally.<sup>15</sup> This means that we do not know whether a meta-ethical view amounts to moral realism until we know whether it fulfils this requirement. Admittedly, there are views about which it is rather obvious that they do not construe moral sentences literally; the view according to which the moral rightness of actions consists in their being desired by some human creature can perhaps serve as an example.<sup>16</sup> But as the meta-ethical literature bears witness, whether a certain view construes moral sentences literally is something that is often quite difficult to know and about which there is much disagreement. In fact, much of the meta-ethical controversy can be understood to concern what the correct literal construal of such sentences consists in. Consequently, it will be quite difficult to know whether a certain meta-ethical position qualifies as moral realism on this conception. This suggests that it is not useful to understand moral realism in this way; in particular, it does not promote clarity in discussions about meta-ethical issues.

---

<sup>11</sup> See e.g. Alston (1996), p. 69; Horwich (1982), pp. 181–182, and Maddy (1990), pp. 6–7, 14.

<sup>12</sup> Cf. Alston (1996), p. 69.

<sup>13</sup> For accounts of moral realism (partly) in terms of ‘face value’ or literal construal, see Railton (1996a), pp. 49–81; Railton (1996b), pp. 51–68; Sayre-McCord (1988 (1986)), pp. 1–23, and Sayre-McCord (1991), pp. 157–158. See also Butchvarov (1988), p. 396. According to Peter Railton, a literal construal of a certain discourse is ‘based upon our ordinary notions, folk theories as influenced by scientific developments, relatively uncontroversial applications, paradigm cases, and so on’ (Railton (1996b), p. 58).

<sup>14</sup> See e.g. Railton (1996a), p. 55–56.

<sup>15</sup> See e.g. Sayre-McCord (1988 (1986)), p. 22.

<sup>16</sup> Cf. Railton (1996a), pp. 67–68, and Sayre-McCord (1988 (1986)), p. 22.

The mentioned difficulty has consequences for how the conception of moral realism under consideration relates to desideratum (i). According to this conception, we do not know whether a meta-ethical view qualifies as moral realism until we know whether it construes moral sentences literally. As a consequence, it is difficult to know to what degree this conception extends to alternative conceptions of moral realism.<sup>17</sup> But it differs at least in one important way from the other conceptions. Suppose that conditions (1) and (2) are fulfilled. Suppose further that there is a meta-ethical view that succeeds to construe some moral sentences literally in the required way. On the present conception, it seems to follow by definition that this view constitutes moral realism. This seems to mean that, granted that the mentioned requirements are fulfilled, moral realism is in a significant respect correct by definition. On other conceptions, it is however a matter of substantial philosophical debate whether moral realism construes moral sentences literally and is the correct view in this respect; it does not follow merely from the assumption that conditions (1) and (2) are fulfilled together with the way in which realism is defined.

*Moral realism in terms of irreducibility*

According to another conception of ontological realist claims, such claims characterise the kind of entities in question as irreducible. Reduction may be understood in terms of a kind of entities, A, being subsumable under a certain kind of entities, B. Accordingly, I will take reduction to imply that a kind of entities A is identical with a set of entities belonging to a certain kind of entities B so that A makes up a subclass of B. For example, mental properties are reduced to physical properties only in so far as they are identical with some physical properties, thus making up a subclass of them. On the present conception of ontological realist claims, such a claim denies that a certain kind of entities is thus reducible.<sup>18</sup> The idea which motivates this conception is, as already suggested, that an ontological realist claim characterises a certain kind of entities as making up a distinct kind of entities in relation to a certain other kind of entities, e.g. that there is something mental separate from the physical. If this notion of ontological realist claims is applied to moral realism, the following condition should be added to (1)

---

<sup>17</sup> However, it extends to the minimalist conception since it implies it.

<sup>18</sup> See e.g. Craig (1998), p. 117; Horwich (1996), pp. 188–190, and Pettit (2002), p. 52.

and (2): (3'') moral properties are irreducible.<sup>19</sup> As I will understand this view, it says that moral properties are irreducible to non-moral properties.

It is possible to argue, I think, that this conception of moral realism is more successful in its fulfilment of the three desiderata than the conceptions considered earlier.

We may start with considering whether it fulfils desideratum (i), that is, whether it extends to alternative conceptions of moral realism.

This conception of moral realism clearly extends to the *minimalist conception*, since it implies it.

Moreover, I think it is reasonable to argue that conceiving of moral realism in terms of irreducibility extends to a reasonable degree the conception which understands moral realism in terms of *mind independence*. On the first-mentioned conception, moral properties are irreducible and hence not identical with any non-moral properties. This implies that moral properties are mind independent in the sense that they are not identical with non-moral properties pertaining to mind. Thus understood, the mind independence of moral properties is a consequence of their being irreducible; since they are not identical with non-moral properties, they are not identical with non-moral properties that pertain to mind.

It might be objected that moral properties being irreducible is compatible with their not being mind independent. To conceive of moral properties as irreducible is compatible with the view that objects have moral properties in virtue of having non-moral properties. Indeed, as I will argue in subsequent chapters, moral realism should be combined with this view. However, this means that the non-moral properties in virtue of which objects have moral properties might consist in non-moral properties pertaining to mind, in which case it might be doubted that moral properties are mind independent.

In reply to this objection, it should be stressed that even if objects have moral properties in virtue of having non-moral properties that pertain to mind, the moral properties *themselves* are mind independent in a certain sense. The reason is, again, that even if objects have moral properties in virtue of having non-moral properties, moral properties are not identical with such properties. The conception according to which moral properties

---

<sup>19</sup> For accounts of moral realism (partly) in terms of irreducibility, see e.g. Arrington (1989), p. 120; Butchvarov (1988), p. 396; Dancy (1998), pp. 534, 536; McNaughton (1998), pp. 41–46, and Platts (1979), pp. 244–245.

are irreducible is therefore in the indicated sense incompatible with the view that such properties are not mind independent.

Furthermore, as was noted above, it is untenable to characterise moral properties as totally mind independent in the sense that they, and the non-moral properties in virtue of which objects have moral properties, are characterised without mentioning of any mental concepts whatsoever. In view of this, it is reasonable to allow that a meta-ethical view which qualifies as an instance of moral realism may claim that objects may have moral properties in virtue of having non-moral properties that pertain to mind.<sup>20</sup> So even if moral realism on the conception under consideration does not guarantee that moral properties are totally mind independent in the indicated way, it is reasonable to believe that it characterises moral properties as mind independent to a reasonable extent.

It can further be argued that conceiving of moral realism in terms of irreducibility extends to a reasonable degree to the conception that characterises this claim in terms of *literal construal* of moral sentences. Even if it can not be presumed that a meta-ethical view which characterises moral properties as irreducible provides a literal construal of moral sentences, such a view is likely to construe moral sentences in a way that does not strike us as obviously non-literal. Since it does not offer a reduction of moral properties, it implies that it is not possible to account for them entirely in terms of non-moral properties. Now, it is presumably accounts of moral properties exclusively in non-moral terms that we are likely to find most informative and hence potentially in conflict with our ordinary notion of them. It is therefore not far-fetched to assume that a meta-ethical view which characterises moral properties as irreducible does not offend our ordinary notion of them, at least not obviously. As a consequence, there is

---

<sup>20</sup> However, should irreducibility be considered too weak to assure mind independence, we might use the notion of mind independence to qualify the basic claim of irreducibility. There are mainly two ways to do this. One way is to state that a necessary condition for a view to qualify as an instance of moral realism is that it denies that the non-moral properties in virtue of which objects have moral properties pertain to mind, where this is conceived in a reasonably strong way. This would then be a fourth condition to add to the three previous ones. However, this alternative has the drawback, noted above, that it is difficult to account for mind independence in a clear and non-arbitrary way. As we saw above, the notion seems to come in degrees. A more attractive alternative is therefore to allow for different moral realist views which vary in strength depending on to what extent they characterise the non-moral properties in virtue of which objects have moral properties as mind independent. However, I will not make use of this strategy.

reason to believe that it does not construe moral sentences which ascribe such properties to objects in a way that we find obviously non-literal.

Let us next consider whether conceiving of moral realism in terms of irreducibility succeeds in relation to desideratum **(ii)**, that is, whether it is an instance of a conception of ontological realist claims that extends to alternative conceptions of such claims.

Earlier we noted that the *minimalist conception* of moral realism does not seem to have any counterparts as ontological realist claims in other fields. In any case, to conceive of ontological realist claims in terms of irreducibility extends to such a conception, since it implies it.

Moreover, I think it is reasonable to argue that conceiving of ontological realist claims in terms of irreducibility extends in an appropriate way to the view which understands such claims in terms of *mind independence*. In doing so, it avoids the difficulties related to this conception. Above I argued that it is mistaken to understand ontological realist claims in terms of mind independence, since such a conception has difficulties in categorising any view that concerns the mental as realist. Furthermore, I argued that it is mistaken to characterise ontological realist claims in terms of independence with respect to any particular phenomenon X, mental or other, since this would make it impossible for there to be such claims about X itself. It should be noticed, however, that these difficulties leave open the possibility to conceive of ontological realist claims in terms of independence, although not in terms of independence of any *particular* phenomenon. I think it can be argued that one way to conceive of ontological realist claims in such a topic neutral manner is in terms of irreducibility. The idea that independence can be understood in terms of irreducibility is close at hand; after all, to hold that a certain entity is irreducible is to imply that it has a certain kind of independence in relation to the kind of entities to which it is irreducible.

According to this understanding of ontological realist claims, the part of such a claim that is concerned with irreducibility consists of two steps.<sup>21</sup> The first step it has in common with other ontological realist claims whereas the second step is allowed to vary depending on the kind of entities that

---

<sup>21</sup> Of course, an ontological realist claim might consist in additional conditions. These conditions may for example say that sentences related to the entities in question have truth-value and that some of them are true in virtue of these entities, i.e. analogues to conditions **(1)** and **(2)** above.

constitutes the object of the particular claim in question. In the *first step* it is claimed about some kind of entities X that it is irreducible. This ‘declaration of independence’ should be understood to be topic neutral in that it does not state that X is irreducible to any particular kind of entities, e.g. mental properties.<sup>22</sup> In the *second step* the ontological realist claim at issue specifies the kind of entities that X is irreducible to, where this is adapted to what is suitable to X.

There are, I think, at least two advantages with characterising ontological realist claims in this way. First, it is, in the first step, topic neutral. This means that it makes possible ontological realist claims as regards a wide range of different kinds of entities. Thus, it does not make the mistake of understanding such claims in terms of independence of a particular phenomenon. Especially, since ontological realist claims on this view are not understood in terms of mind independence, but in terms of topic neutral independence *qua* irreducibility, there are such claims also about mental properties. Hence, it does not have the difficulty related to characterising ontological realist claims in terms of mind independence noted above. Second, this conception of ontological realist claims allows, in the second step, ontological realist claims to vary concerning the kind of entities of which independence is claimed in a way that is tailor-made to these entities. As a consequence, it makes possible ontological realist claims about different kinds of entities in a way that agrees with our views of what realist claims about these entities should amount to. For example, it makes possible an ontological realist claim according to which a certain kind of entities, e.g. scientific entities, is irreducible to entities that pertain to mind. This means that in the cases it is appropriate, the proposed characterisation of ontological realist claims extends in a quite substantial way to the view that such claims should be understood in terms of mind independence.

The following examples might illustrate how this conception of ontological realist claims works. Ontological realist claims about scientific

---

<sup>22</sup> It might be trivially true that a certain kind of entities X is irreducible to *some* kind of entities. But the first step might simply be understood as the view that X is irreducible to a definite kind of entities in a non-trivial way, but that it at this stage is not decided what this kind of entities consists in. Alternatively, it might be understood as the claim that X is irreducible to the kind of entities that constitutes its ‘complement’, whatever kind of entities that is. In the case of mental properties, the ‘complement’ is constituted by physical properties; in the case of scientific entities, it is presumably constituted by entities that pertain to mind, etc.



entities can be interpreted to embrace the topic neutral irreducibility claim about scientific entities (first step) and then to assert that such entities are irreducible to properties pertaining to mental states, understood in a rather broad sense, including beliefs and observations (second step). Operationalism, for example, does not fulfil this condition, and it would consequently not qualify as an ontological realist claim on this conception. Ontological realist claims about mental properties might be interpreted to embrace the topic neutral irreducibility claim about such properties (first step) and then to assert that mental properties are irreducible to physical properties, understood in a rather broad sense (second step). Since for example Cartesian dualism implies that mental properties are irreducible to physical properties, it would qualify as an ontological realist claim on this conception. However, some other views, such as behaviourism, do not fulfil this criterion and would consequently not qualify as realist. Moral realism might be interpreted to embrace the topic neutral irreducibility claim about moral properties (first step) and then to assert that such properties are irreducible to non-moral properties (second step).<sup>23</sup>

It should be stressed that this proposal is not intended as a description of what authors mean when they label a certain view ‘realism’. Rather, it is intended as a way of understanding ontological realist claims that lends a certain coherence to various assertions in relation to them. Needless to say, here I am only able to give a very short outline of this proposal.

To conceive of ontological realist claims in terms of irreducibility also extends in a reasonable way to the conception that understands such claims in terms of ‘*face value*’. The reason for this being so should be familiar from what was said above. A view which states that a certain kind of entities is not reducible to entities belonging to another kind implies that the nature of the former kind of entities cannot be characterised completely in terms of the latter. However, it is presumably an account that does this that we are likely to find most informative and hence as characterising the entities in a way that conflicts with our ordinary notion of them. It is therefore reasonable to assume that an account of these entities which characterises them as irreducible does not characterise them in a strikingly unfamiliar way and hence not in a way that is obviously incompatible with our ordinary notion of them.

---

<sup>23</sup> In the discussion to follow, I will not make use of this distinction.

Lastly, let us consider whether the conception of moral realism in terms of irreducibility fulfils desideratum (iii), that is, whether it is serviceable in meta-ethical discussions. I think it is. Whether a meta-ethical view is an instance of moral realism is according to this conception a rather straightforward matter, at least in principle; it turns on whether the view at issue reduces moral properties to non-moral properties. Since moral realism on this conception is a matter of ‘either or’, this conception promotes clarity in discussions about meta-ethical issues.

#### *4. Moral Realism and Its Opponents*

In the last section, I argued that we should adopt a conception of moral realism according to which this view comprises three conditions: first, moral sentences have truth-value; second, there exist moral properties so that some of these sentences are true, and, third, moral properties are irreducible to non-moral properties. These claims constitute, I think, necessary and jointly sufficient conditions for a meta-ethical view to be an instance of moral realism.

In this section, I will examine what the meta-ethical views involved in these three claims amount to and contrast them with the meta-ethical views that are their respective opposites.

##### *Cognitivism contra non-cognitivism*

Cognitivism may roughly be understood as the view that moral sentences have truth-value; that is, that they are true or false. Hence, the first condition of moral realism is an affirmation of cognitivism. Non-cognitivism may roughly be understood as the view that moral sentences lack truth-value.

What cognitivism and non-cognitivism amount to can be seen more exactly if we consider what they say about a certain type of sentences. Many sentences involve terms in virtue of which properties are ascribed to objects. Such sentences are true if the objects have the properties and false if they do not. By appearance, many moral sentences belong to this type or are reasonable interpreted as belonging to it. On cognitivism, this impression is correct because such moral sentences ascribe properties to objects in just the same way as other sentences of this type. That is, such moral sentences contain moral terms in virtue of which moral properties are ascribed to

objects, and if the objects in question have the moral properties, the sentences are true, false if they do not. Non-cognitivism denies this and claims that, despite appearances, moral sentences differ from other sentences of a similar form. On this view, moral sentences of this type do not contain moral terms in virtue of which moral properties are ascribed to objects. Hence, non-cognitivism denies that moral sentences are true or false in the indicated way. Since non-cognitivism implies that no moral sentences are thus true, I will take it to imply that there is nothing that can make such sentences true in the indicated way; that is, I will take it to imply that there are no moral properties.<sup>24</sup>

In view of this difference, cognitivism and non-cognitivism also differ as to what kind of states moral sentences are used to express. It might be claimed that on both views such sentences are used to express moral judgements; the term ‘moral judgement’ may then be taken to be neutral between cognitivism and non-cognitivism. However, they disagree as to what a moral judgement consists in. According to cognitivism, it consists, at least primarily, in a belief. Thus, on this view, moral sentences of the kind mentioned express beliefs to the effect that objects have certain moral properties. According to non-cognitivism, a moral judgement does not primarily consist in a belief but in a non-cognitive state of some sort. Thus, on this view, moral sentences express such states.

It might be asked what view on truth that cognitivism and non-cognitivism involve. As far as I understand, there is no need for advocates of these positions to commit themselves to any particular view on this notion. These meta-ethical positions should thus, *mutatis mutandis*, be possible to combine with various views on truth. However, it is important that the view on truth involved in cognitivism and non-cognitivism is not different from the one involved in non-moral areas. As Geoffrey Sayre-McCord points out, the view at issue should be ‘seamless’ in the sense that ‘whichever theories of meaning and truth are offered for the disputed claims must be extended as well to apply to all claims’.<sup>25</sup> Above all, it should not

---

<sup>24</sup> Cf. Shafer-Landau (2003), p. 18, and Thomson in Harman and Thomson (1996), p. 96.

<sup>25</sup> Sayre-McCord (1988 (1986)), p. 6. Cf. Pettit (2001), pp. 242–245.

be weaker than the view on truth appealed to in other areas. This requirement finds support in desideratum (ii).<sup>26</sup>

It should be mentioned that according to some non-cognitivists, moral sentences can express beliefs, although merely secondarily. On this view, moral terms have two kinds of senses: a primary sense, which consists in a certain non-cognitive attitude, and a secondary sense, which consists in certain non-moral properties.<sup>27</sup> These properties can be understood to constitute the speaker's moral criteria or principles that determine towards which objects she has the non-cognitive attitude and hence to which objects she applies the term. As a consequence of the secondary sense of moral terms, moral sentences can be taken to express beliefs about these non-moral properties. The reason to call the former 'primary sense' and the latter 'secondary sense' might perhaps be understood in the following way. Suppose a person asserts a sentence in which a moral term is involved, such as 'It's right to give money to the Red Cross'. When 'right' figures in such a sentence, it is necessarily used to express a non-cognitive attitude in the sense that unless a speaker has the attitude in question, her use of the term is

---

<sup>26</sup> It might be argued that non-cognitivists can claim that moral sentences have truth-value. They can do this because they might adopt a deflationist or minimalist conception of truth. For example, they may maintain that to say that a sentence is true merely is to affirm it. This possibility is especially exploited by quasi-realists. (See e.g. Blackburn (1993 (1988)), pp. 166–181, and Blackburn (1998), pp. 68–83.) However, advocates of this view are non-cognitivists according to the characterisation above. The main reason is that they deny that moral sentences involve terms in virtue of which these sentences ascribe properties to objects in the same way as other, non-moral, sentences of a similar form. They deny therefore that any moral sentences are true or false in virtue of objects having or lacking these properties. Accordingly, they deny that there are any moral properties that can make sentences true in that way. Moreover, they deny that moral sentences express beliefs, at least primarily. In what follows, I will continue to say that non-cognitivists claim that moral sentences lack truth-value. What I will have in mind is that they claim that moral sentences are not true or false in the qualified sense described above.

It might be objected to this reasoning that non-cognitivists can claim that moral sentences ascribe moral properties and express beliefs. Non-cognitivists can claim this because just as they may adopt a deflationist or minimalist conception of truth, they can adopt such a conception of properties and beliefs. For example, they might argue that to say that an object has a moral property merely is to affirm a sentence to this effect or to say that a person has a moral belief merely is to say that she is disposed to affirm a certain sentence. However, advocates of this view would still be classified as non-cognitivists according to my characterisation. The reason is that they deny that moral sentences ascribe properties to objects and express beliefs in the same sense as other, non-moral, sentences do so. On their view, moral properties and beliefs are not full-blown properties and beliefs, but quite different from properties and beliefs in other areas. Hence, cognitivism can be understood as a thesis which stresses the continuity between moral discourse and ordinary discourse.

<sup>27</sup> See e.g. Hare (1952), pp. 118–126, and Hare (1997), pp. 48–56.

incorrect. By contrast, the term is not necessarily used to claim that objects have any particular non-moral properties in this way. The non-moral properties associated with 'right' may consequently vary with different speakers even if their use of the term is correct. This means that two persons use 'right' in the same meaning only if they use it in the same primary sense. However, they may use the term in the same meaning even if they differ as regards the secondary sense they use it to have. As we will see in chapter 7, cognitivists can in a like manner argue that moral sentences express non-cognitive states, although secondarily.

#### *Success-theory contra error-theory*

According to success-theory, there exist moral properties that make some moral sentences true.<sup>28</sup> Hence, the first condition of moral realism is an affirmation of success-theory. As we saw above, cognitivism claims that certain moral sentences contain moral terms in virtue of which moral properties are ascribed to objects. Advocates of success-theory agree and add the claim that moral terms refer to existing moral properties with the consequence that some moral sentences of this kind are true. Advocates of error-theory also agree with cognitivism, but deny the second claim. On this view, there are no moral properties, with the consequence that no moral sentences of the mentioned kind are true; to the contrary, all such sentences are false.<sup>29</sup>

It might be asked what notion of properties is involved in success-theory and error-theory. In analogy with the answer to the corresponding question as regards truth, I think the proper response is that these views need not be committed to any particular view in this regard. Thus, success-theory and error theory should, *mutatis mutandis*, be possible to combine with various views on properties. It is however important that the notion of property involved in these views is not different from the one appealed to in other areas. The reason is basically the same as the one regarding truth.

---

<sup>28</sup> I have borrowed the term 'success-theory' from Sayre-McCord (1988 (1986)), p. 10.

<sup>29</sup> It might be argued that some moral sentences actually are true on error-theory. One example might be 'Everything is morally permissible'. However, I focus on sentences that attribute moral properties as rightness, wrongness, goodness and badness to objects. As I will understand error-theorists, they deny that there are any true moral sentences of these kinds.

### *Realism contra Reductionism*

Moral realism is the affirmation of cognitivism and success-theory together with the claim that moral properties are irreducible to non-moral properties. Moral reductionism is the denial of this third claim. Moral reductionists accept cognitivism and success-theory but maintain that moral properties *are* reducible to non-moral properties. Henceforth I will refer to these views simply as ‘realism’ and ‘reductionism’, respectively.

We may now inquire what is involved in these views. What is it realists deny and reductionists affirm?

We may start by considering the scope of these views. Earlier I said, somewhat vaguely, that realists deny, whereas reductionists affirm, that moral properties are reducible to non-moral properties. The moral properties I have in mind are ‘thin’ moral properties, particularly moral rightness, wrongness, goodness and badness.<sup>30</sup> In what follows, I will at times represent such a moral property as ‘M’.

Let us continue by considering what is involved in reductions of moral properties. The basic component of such a reduction is, as was indicated in the last section, the claim that moral properties are *identical* with certain non-moral properties.<sup>31</sup> (It might appear paradoxical to say that moral properties are identical with non-moral properties. However, as non-moral properties are characterised below, I think this impression is mistaken. But if this usage is thought to be peculiar in spite of that characterisation, we might label these properties in some other way.) This requirement is motivated by the following considerations.

First, although reduction is a notion that has been debated in philosophy, it seems nowadays fairly uncontroversial to claim that a necessary condition for a kind of entities A to be reduced to a certain kind of entities B is that A is identified with a set of entities belonging to kind B,

---

<sup>30</sup> A weaker form of realism is had if it is claimed that only some of these ‘thin’ moral properties are irreducible. In chapters 6 and 7, I will focus on rightness. This means that even though I take realists to argue that all the mentioned properties are irreducible, they may make a stronger case in relation to this particular moral property. As I consider moral properties unless I indicate otherwise, I will write ‘right’ instead of ‘morally right’, etc.

<sup>31</sup> For an account of the importance of identity claims in reductions, see e.g. Hooker (1981), pp. 201–236.

so that A makes up a subclass of B. It is, in any case, uncontroversial in meta-ethics.

Second, as was indicated in previous sections, lack of identity is required for the realist claim that moral properties make up a distinct kind of properties; unless moral properties are non-identical with non-moral properties, they do not constitute a kind of properties distinct from non-moral properties.

Correspondingly, the requirement that reductions involve identity claims is tightly connected to various reasons for reductions.<sup>32</sup> Here I will only mention some reasons for reducing moral properties, but they have counterparts in other fields of philosophy. One reason is to simplify ontology. It might be argued that ontology can be simplified by subsuming moral properties under the category of non-moral properties. This requires that moral properties are identical with non-moral properties. Another reason for reduction is to save the existence of moral properties. Sometimes it is suspected that moral properties would be for example metaphysically or epistemologically deviant and that their existence therefore should be denied. In order to uphold the existence of moral properties, it might be argued that they are reducible to non-moral properties that are not deviant in these respects. This enterprise also seems to require that moral properties are identical with certain non-moral properties. A third reason for reduction is to be able to explain certain characteristics connected to moral properties. It might for example be argued that if moral properties are not reduced to non-moral properties, it is not possible to account for how they can be involved in causal explanations of various facts. This enterprise also seems to require that moral properties are identical with certain non-moral properties.

Realists, on the other hand, argue that these reasons for reductions are misguided. They may argue that ontology cannot be simplified with respect to moral properties because such properties actually make up a separate category of properties. And they may argue that moral properties are neither metaphysically nor epistemologically deviant even if they are not identified with non-moral properties. Further, they may argue that the characteristics connected to moral properties can be explained, and perhaps even *better*

---

<sup>32</sup> For various reasons for reductions, see e.g. Alston (1996), pp. 69–73; Hooker (1981), pp. 201–204, and Mumford (1998), pp. 185–190. For a discussion of reduction in ethics, see e.g. Railton (1989), pp. 155–163, and Railton (1993b), pp. 317–320.

explained, without assuming identity with non-moral properties. In subsequent chapters we will find examples of these ways of arguing.

Reduction is an asymmetric notion since it implies that one kind of entities, the entities to which reduction is made, have some sort of priority over the entities that are reduced. Accordingly, although we may want to say that moral properties are reduced to non-moral properties, we are not prepared to say that the converse holds. However, identity is a symmetric notion. In consideration of this, it can be argued that even if identity constitutes a necessary condition for reduction, it does not constitute a sufficient condition. On this view, a reduction has to involve at least one additional component to become asymmetric.<sup>33</sup> As a consequence, this view raises the question of how the additional asymmetric component of a reduction should be understood. This is a difficult issue that troubles philosophers in different fields, and I cannot deal with it here. However, I think it can be argued that identity claims are closely associated with certain asymmetric components. Most fundamentally, appropriate identity claims imply a certain type of ontological asymmetry. Above it was suggested that a necessary condition for moral properties to be reduced to non-moral properties is that the former kind of properties is identified with a set of the latter kind of properties, so that the former make up a subclass of the latter. This means that in so far as moral properties are reduced to non-moral properties, they are subsumed under a more inclusive class of properties. In this case, the asymmetric component consists in a relation holding between moral properties and the more inclusive class of non-moral properties of which they are a subclass. However, the examples of reasons to reduce moral properties above suggest other ways in which identity claims might be connected to asymmetric components. As regards the first reason for reduction, the identification of moral properties with certain non-moral properties might mean that they are more fundamental or simple than one believes. As regards the second reason for reduction, the identification of moral properties with certain non-moral properties might mean that they are less philosophically troublesome than one believes. As regards the third reason for reduction, the identification of moral properties with certain non-moral properties might mean that the characteristics connected to

---

<sup>33</sup> For a discussion of asymmetry in reductions, see Mumford (1998), pp. 172–176, 183–191.



moral properties are more readily explainable than one believes. In these cases, the asymmetric relation holds between, on the one hand, moral properties as one might believe they are and, on the other hand, how they are on the assumption that they are identified with certain non-moral properties. Unfortunately, I am unable to account for how the asymmetric component indicated by these claims should be explained more exactly. However, whereas the first-mentioned asymmetric component is ontological, the latter three are presumably epistemological or explanatory.

Since a necessary condition for reduction is identity between properties, it becomes relevant to ask what constitutes property identity. In the literature there are different views about this notion, but at least one necessary condition seems uncontroversial. It is generally agreed that a property A is identical with a property B only if it holds necessarily that if something has A it has B and vice versa. That is, A and B are identical only if they are necessarily co-extensive. Suppose a reductionist claims that a moral property M is identical with a non-moral property G. According to the mentioned requirement on identity, she is then committed to the truth of a necessary biconditional:

It is necessary that, for any object x, x is M if and only if x is G.

Reductionists might maintain that G is a complex property, i.e. that G is a non-moral property which in turn consists of non-moral properties. In case G is claimed to be complex, I will consider it to consist of a conjunction of non-moral properties. However, I will not allow that G consists of a disjunction of non-moral properties. Although it is not completely uncontroversial whether conjunction is a proper way of forming properties, it is, as far as I understand, fairly generally accepted that it can have this function. However, it is highly controversial whether disjunction is a proper way of forming properties. If disjunction cannot form properties, the resulting necessary biconditional would not provide a basis for an identity claim. It might also be thought that if disjunctions are allowed to form properties, there is a powerful argument against realism. I will consider this argument in chapter 8.<sup>34</sup>

As I have formulated realism and reductionism, they deny and affirm, respectively, that moral properties are reducible to non-moral properties.

---

<sup>34</sup> In chapter 5, we will see that there might be a relativist version of reductionism. This view might be understood to imply that more than one necessary biconditional of the indicated kind holds.

This raises the question of how the distinction between these two kinds of properties should be understood.

It should first be emphasised that as I understand this distinction, identity between them is not ruled out beforehand, since I do not use the terms ‘moral properties’ and ‘non-moral properties’ in such a way that a property is a moral property only if it is not a non-moral property and vice versa.<sup>35</sup> Rather, I conceive of moral properties and non-moral properties as two kinds of properties about which it is an open question whether moral properties make up a subclass of non-moral properties. That is, as far as the distinction between these two kinds of properties goes, moral properties may constitute a separate kind of properties (as realists claim), or a subclass of non-moral properties (as reductionists claim).

Many meta-ethicists appeal to moral and non-moral properties to characterise various meta-ethical views. They often leave the distinction uncommented, thus suggesting that it is unproblematic. However, R. M. Hare tries to clarify this distinction.<sup>36</sup> Mapped over to the terminology adopted here, Hare’s account of non-moral properties might be formulated in the following way: a property is a non-moral property in so far as it can be completely characterised without the use of any moral terms.<sup>37</sup>

We can employ this account of non-moral properties to illuminate the difference between reductionism and realism. According to reductionism, a moral property M is identical with a non-moral property, i.e. a property which is such that it can be completely characterised without the use of any moral terms. This view implies that there is a necessary biconditional with ‘x is M’ on the left-hand side and on the right-hand side a property G for which this holds. Realism, on the other hand, denies that a

---

<sup>35</sup> Cf. Baldwin (1985), p. 32.

<sup>36</sup> Hare (1997), p. 64. However, similar accounts are found in Pigden (1991), p. 421; Railton (1989), p. 160; Sayre-McCord (1997a), p. 281, and Timmons (1999), p. 48.

<sup>37</sup> This account should presumably be qualified. It seems trivially true that a property can be completely characterised without the use of any moral terms. For example, rightness might be characterised as the property people think of at a certain moment or as the property about which they make certain utterances. And this holds presumably even if it is irreducible. To avoid this problem, we may put some restrictions on the mentioned characterisation. The two examples suggest that at least two such restrictions are needed. First, the property should not be characterised as the object of an attitude. Second, it should not be characterised as something about which an utterance is made. However, other qualifications might be required as well.

moral property M is identical with a property which is such that it can be characterised without the use of any moral terms.<sup>38</sup>

I previously mentioned that reductionism states that moral properties are identical with certain non-moral properties, so that the former kind of properties makes up a subclass of the latter kind. The account of non-moral properties can be employed to elucidate this claim. If, as reductionism says, moral properties are identical with certain non-moral properties, and hence constitute a subclass of non-moral properties, they can quite naturally be completely characterised without the use of any moral terms. On the other hand, if, as realism says, moral properties are not identical with any non-moral properties, and hence constitute a class of properties separate from the class of non-moral properties, they cannot be completely characterised without the use of any moral terms.

However, it should be admitted that this account is not very helpful. For one thing, it is rather vague. Moreover, to characterise non-moral properties by means of moral terms raises the question of how moral properties should be understood. It seems difficult to account for them without making use of moral terms. It might then be suspected that the distinction between moral and non-moral properties cannot be accounted for in a way that is not circular.

Unfortunately, I do not know how to spell out the distinction in a satisfactory way. Above all, I do not know how to account for it without circularity. However, two comments might serve to alleviate this worry. First, the distinction between moral and non-moral properties appears to a fair extent to be intuitively clear. The fact that most authors do not dwell on it, thereby suggesting that it is unproblematic, is a sign of this. Another sign is that we normally do not seem to have any troubles classifying properties as moral or non-moral, even though there might be exceptions. This holds especially as regards the kind of moral properties I will be concerned with, i.e. 'thin' moral properties, and non-moral properties. For example, we have no problems judging that the term 'right' in certain contexts refers to moral rightness, and hence to a moral property, whereas e.g. 'maximising happiness' refers to a non-moral property. Moreover, having found that the terms refer to a moral property and a non-moral property, respectively, does not prevent us from leaving it open that these terms refer to the same

---

<sup>38</sup> Cf. Hare (1997), pp. 64, 82.

property. Second, a general distinction between the moral and the non-moral is presupposed in a number of considerations, both in philosophical ethics and in everyday situations. We may for example ask whether the moral and the non-moral conflict or at certain points converge. The distinction between moral and non-moral properties is part of this general way to conceive of matters. This suggests that in so far as the distinction between moral and non-moral properties is problematic, this is part of a larger difficulty that concerns the general relation between the moral and the non-moral. It would thus not be a difficulty exclusive for the present work or for meta-ethics in general.

### *Realism and naturalism*

The way I understand realism may be contrasted with a more traditional conception of what this position amounts to, influenced by the work of G. E. Moore.<sup>39</sup> It takes, like my conception, its point of departure in the idea that realism characterises moral properties as making up a separate kind of properties. It therefore states that the characteristic feature of realism is the claim that such properties are irreducible. However, on this conception realism does not say that moral properties are irreducible to non-moral properties, but that they are irreducible to *natural properties*.<sup>40</sup>

The main difficulty with this alternative conception of realism is that it rules out as non-realist a meta-ethical view that reasonably qualifies as realist. The view I have in mind says that moral properties are irreducible to non-moral properties but that they nevertheless are natural properties. On this view, moral properties constitute a separate category of properties, since they are distinct from the category of non-moral properties. Still, they are members of the category of natural properties. Suppose, as both my conception and the alternative one suggest, that the image realism is getting at is that moral properties constitute a separate kind of properties and that the characteristic feature of realism is therewith that moral properties are irreducible. Given this governing idea, it is difficult to see why the view under consideration should not be allowed as an instance of realism since it, after all, states that moral properties make up a separate kind of properties.

---

<sup>39</sup> Moore famously understood naturalism as the view that goodness is reducible to natural properties; see e.g. Moore (1993 (1903)), pp. 91–93.

<sup>40</sup> On this conception, realists also claim that moral properties are not reducible to supernatural properties.

This holds particularly since this view is one that has been found attractive by some influential meta-ethicists widely accepted as realists, most prominently Nicholas L. Sturgeon and David O. Brink.<sup>41</sup> Moreover, as is implied by the discussion in the last section, the view that moral properties are irreducible but nonetheless natural qualifies as realist in terms of the desiderata of a successful conception of realism.

These considerations suggest that there is a distinction between two kinds of naturalism.<sup>42</sup> According to *reductive naturalism*, a moral property M is reduced to, and therefore identical with, a non-moral property that is a natural property. On this view, a necessary biconditional of the kind mentioned above holds where a non-moral property G consists in a natural property. According to *non-reductive naturalism*, M is a natural property, but it is not reduced to any non-moral property and hence not to any non-moral property that consists in a natural property. Realism is incompatible with reductive naturalism but compatible with non-reductive naturalism. Of course, even if realists may hold that moral properties are natural properties, they need not. They may instead believe that moral properties are supernatural properties, or they may be neutral in this respect.

These remarks raise the question of how natural properties should be characterised. This is a difficult issue that has troubled meta-ethicists ever since Moore made use of this notion—not the least it troubled Moore himself. As far as I understand, there are basically three conceptions of natural properties in the literature that may be relevant to realism.<sup>43</sup> According to Moore's official characterisation, natural properties are properties that are the subject matter of science.<sup>44</sup> According to a second conception, natural properties are properties that are known empirically.<sup>45</sup> According to a third conception, natural properties are properties that are

---

<sup>41</sup> See e.g. Brink (1989), pp. 176–177; Sturgeon (1984), pp. 59–60, and Sturgeon (2003), pp. 536–540, 553.

<sup>42</sup> Cf. Brink (1989), p. 9, and Shafer-Landau (2003), p. 63. A similar distinction can be drawn between reductive supernaturalism and non-reductive supernaturalism by exchanging natural properties for supernatural properties.

<sup>43</sup> For an overview of the various ways in which the notion of natural properties has been understood, see Copp (2003), pp. 182–185. For overviews of Moore's various attempts to define natural properties, see Baldwin (1985), pp. 25–30, and Sturgeon (2003), pp. 541–556.

<sup>44</sup> Moore (1993 (1903)), pp. 91–92. See also e.g. Brink (1989), pp. 9, 157; Shafer-Landau (2003), p. 59, and Timmons (1999), p. 13.

<sup>45</sup> See e.g. Cargile (1989), pp. 142–144; Copp (2003), pp. 185–187, and Railton (1989), p. 154.

part of nature in the sense of being involved in its causal relations.<sup>46</sup> All these conceptions can of course be developed in different ways. In chapter 9, I will argue that realists can defend the view that moral properties are involved in causal explanations, in which case moral properties are natural according to the third conception of such properties.

##### 5. *Some Arguments against Non-cognitivism and Error-theory*

In this section, I will very briefly consider some difficulties with two of the meta-ethical positions mentioned above, non-cognitivism and error-theory, which provide at least *prima facie* reasons not to adopt these views. These difficulties, most of them quite well known, justify that we turn to the two other meta-ethical positions, reductionism and realism, and consider whether any of them are more plausible. Now, there are important arguments that are thought to speak in favour of non-cognitivism and error-theory. In chapter 5, 7 and 8, I will discuss the most important meta-ethical issues considerations of which have been taken to provide support to these views. It seems reasonable to assume that if these issues can be accounted for by reductionism or realism, this, together with the difficulties to be considered below, suggests that non-cognitivism and error-theory can be rejected. However, if both reductionism and realism face problems as difficult as those troubling non-cognitivism or error-theory, or if both fail to explain the mentioned issues, we would have reasons to reconsider non-cognitivism and error-theory. Fortunately, I do not think this will turn out to be necessary.

Consider first non-cognitivism.<sup>47</sup> The most important difficulties with this view are centred on the fact that there are a number of indications that moral sentences have truth-value in the way indicated above. First, many moral sentences seem to involve terms in virtue of which they ascribe moral properties to objects. Hence, they seem to be true or false just in the way other sentences of this form are. Non-cognitivists, denying that moral sentences have truth-value, have thus to understand such sentences in a way that is not consistent with their appearance. Moreover, some of our

---

<sup>46</sup> See e.g. Baldwin (1985), pp. 34–35; Smith (1999), pp. 93–101, and Sturgeon (2003), p. 538.

<sup>47</sup> For overviews of difficulties with non-cognitivism, see e.g. Brink (2000), pp. 196–200; Darwall, Gibbard and Railton (1992), pp. 144–152; Miller (2003), esp. pp. 38–51; Shafer-Landau (2003), pp. 22–38, and Thomson in Harman and Thomson (1996), chap. 7. Of course, non-cognitivism has more difficulties than those I mentioned here.

epistemic attitudes in relation to moral matters are such that moral sentences have to have truth-value. For example, we may take ourselves to *know* that Hitler was bad, to *believe* that it is right to let women decide whether they should have an abortion, to be *unsure* whether it is right to give money to beggars, to have been *mistaken* in believing that homosexual behaviour is wrong, etc. Such epistemic attitudes seem to require that moral sentences can be true or false. Since non-cognitivists deny this, they have difficulties accounting for these epistemic attitudes.

Relatedly, our notion of moral disagreement requires that moral sentences have truth-value. The most obvious indication of this is presumably that, at least normally, we consider a person who asserts a moral sentence to be genuinely disagreeing with a person who denies this sentence in such a way that their sentences are contradictory. If one person asserts a moral sentence and another denies it, we take it that they hold views that are incompatible. Thus, we take it that it is impossible that both are right; rather, one is right and the other wrong. Since non-cognitivists deny that moral sentences have truth-value, they seem to have difficulties accounting for moral disagreement.

There is one indication of moral sentences being truth-evaluative that appears even more troubling for non-cognitivism. On the received view, for sentences to function as premises and conclusions of valid inferences, they have to have truth-value. Now, it seems evident that moral sentences are involved in valid inferences. For example, we can validly infer from the premises 'If it is wrong to beat one's own children, it is wrong to beat other people's children', and 'It is wrong to beat one's own children' the conclusion 'It is wrong to beat other people's children'. Since non-cognitivists hold that moral sentences do not have truth-value, they are challenged to explain how such inferences can be valid.

A related difficulty emerges in relation to the view that moral terms are used to express non-cognitive attitudes. This view has perhaps some plausibility in contexts where moral sentences are asserted, e.g. 'It's wrong to beat one's own children'. However, it seems difficult to maintain this view when the sentences are unasserted. For example, if a person asserts the sentence 'If it's wrong to beat one's own children, it's wrong to beat other people's children', she does not assert a sentence to the effect that it is wrong to beat one's own children. It seems therefore implausible to hold

that the term ‘wrong’ in such a context is used to express a non-cognitive attitude. The upshot is that non-cognitivists have to claim that moral terms have different meanings depending on the context in which they figure. However, we take moral terms to have the same meaning independently of context. Moreover, in order for moral sentences to be part of valid inferences, the terms involved in these sentences have to have the same meaning throughout such inferences. The mentioned problem thus underlines the difficulty of non-cognitivism to account for moral sentences being involved in valid inferences.

Consider next error-theory.<sup>48</sup> According to this view, moral sentences do have truth-value. Error-theory is consequently not vulnerable to the same criticism as non-cognitivism. For example, on this view moral sentences can be involved in valid inferences. However, in some respects error-theory has difficulties similar to those of non-cognitivism. One such difficulty concerns our epistemic attitudes in relation to moral matters. As mentioned above, we take ourselves to know e.g. that Hitler was morally bad. Since error-theorists claim that no moral sentences in which moral properties are ascribed to objects are true, they have to say that we do not have any such knowledge. Similar considerations may hold in relation to some of the other attitudes mentioned above.

Another difficulty with error-theory concerns moral disagreement. On this view, in contrast to non-cognitivism, people involved in moral debates may genuinely disagree, since they may assert contradictory sentences. However, in one important respect, error-theory conflicts with our notion of moral disagreement. Suppose someone asserts a sentence that ascribes a moral property to an object, e.g. ‘That action is right’. Suppose someone else denies this sentence and claims that it is not the case that the action in question is right. According to error-theory, since there are no moral properties, all sentences that ascribe moral properties to objects are false. Hence, a person who denies such a sentence is right and her opponent wrong.<sup>49</sup> As I indicated above, we seem to believe that in moral disagreements, someone is right and the other wrong. However, error-theory gives an account of why this holds that opposes our conception of moral disagreement. Suppose a person who denies a sentence that ascribes a

---

<sup>48</sup> Error-theory has not been criticised to the same extent as non-cognitivism. But see e.g. Pettit (2001), pp. 255–258, and Wright (1992), chap. 1.

<sup>49</sup> Cf. Pettit (2001), p. 241.



moral property to an object, such as the one just mentioned, is right. According to our notion of moral disagreement, she is right because as regards the particular moral issue under consideration she has identified the correct position to take. In terms of the examples, she has recognised that the action in question is not right, in contrast to certain other actions that have this property. However, according to error-theory, a person who holds that a sentence which ascribes a moral property to an object is false will *always* be right. Moreover, she will be so for the simple reason that there are no moral properties that can make such sentences true and not because she has identified the correct position to take regarding the moral issue under debate. Hence, error-theory gives a wrong reason for why she is right.

Moreover, in at least one respect error-theory seems to be less plausible than non-cognitivism. According to non-cognitivism, moral sentences do not have the function to ascribe moral properties to objects but to give expression to non-cognitive states. Above we have seen some reason to believe that this view is mistaken. However, according to non-cognitivism, moral sentences are at least able to fulfil the function they are claimed to have on this view. According to error-theory, some moral sentences do have the function to ascribe moral properties to objects, but since there are no moral properties they can never fulfil it. In fact, with this view we are under the influence of a gigantic and collective misunderstanding about what we are doing when we use moral language. This seems *prima facie* untenable. Surely, we do not take moral sentences to systematically miss their target in this flagrant way. To the contrary, we use such sentences in the assurance that they do what we use them to do, however that should be understood. Of course, this does not prove that error-theory is false, but it suggests that we should try to find an alternative for it.

As is very well known, many of the objections to non-cognitivism and error-theory considered above have been replied to by advocates of these views. In the case of non-cognitivism, some of these responses have given rise to quite sophisticated versions of this view, which means that my characterisation of this view above is oversimplified. However, as already mentioned, I will not consider these responses. They would be necessary to consider if there were convincing arguments in favour of non-cognitivism or error-theory which cannot be accounted for by any other view, or if the

alternative views have at least as serious difficulties of their own. However, since I will argue that this is not the case, I will not do so.

## 6. *Outline of the Argument*

In the last section, we saw that there are some arguments against error-theory and non-cognitivism which justify considering whether any of the two other meta-ethical positions, reductionism and realism, are more plausible.

In the next two chapters, I examine what I consider to be the two main forms of reductionism: analytic reductionism and synthetic reductionism. In *chapter 2*, I argue that an amended version of G. E. Moore's open question argument suggests that analytic reductionism is mistaken. In *chapter 3*, I argue that considerations of two questions in relation to certain thought experiments suggest the same conclusion regarding synthetic reductionism. However, the arguments of these two chapters are tentative and there are consequently reasons to return to reductionism later on in the investigation.

Some meta-ethicists would argue that the failure of reductionism means that either non-cognitivism or error-theory has to be correct. In *chapter 4*, I argue that this conclusion can be avoided by adopting realism. Now, even if realists claim that moral properties are irreducible, they have to account for our notion that objects have moral properties in virtue of having non-moral properties. In this chapter, I maintain that realists can do so by employing what I call 'the realist formula'. The realist formula acquires confirmation from the discussion of analytic and synthetic reductionism in chapters 2 and 3 by explaining our responses to the questions discussed there. Moreover, since the realist formula implies that moral properties are irreducible to non-moral properties, these responses provide support to realism.

Above I mentioned that although there are reasons to believe that non-cognitivism and error-theory are mistaken, there are also important arguments that have been proposed in support of these views. This means that non-cognitivism and error-theory cannot be rejected until these arguments have been responded to. I also mentioned that although the arguments proposed in chapters 2 and 3 suggest that reductionism is mistaken, this view is in need of further examination. In chapters 5, 6 and 7,

I argue that realism is able to explain certain meta-ethical issues concerning moral disagreement, moral reason and moral motivation and that it therefore is reasonable to maintain that realism is preferable to non-cognitivism, error-theory and reductionism. With regard to non-cognitivism and error-theory, I argue that realism is able to explain some of the issues which are argued to support these views. In combination with the difficulties with non-cognitivism and error-theory mentioned in the last section, this provides reasons to believe that realism is preferable to these two views. Moreover, I argue that realism is able to explain certain meta-ethical issues better than reductionism, which gives reasons to believe that realism is preferable to this view as well. In particular, I argue that realism is more successful than reductionism in accounting for the meta-ethical issues that are taken to support non-cognitivism or error-theory. Accordingly, realism is, in contrast to reductionism, capable of defending the view that moral sentences have truth-value and that there are moral properties. However, it might be objected that realism suffers from difficulties of its own which mean that even if it is preferable in the mentioned respects, it should not be adopted. I will respond to some of the most important arguments to this effect in chapters 8 and 9.

In *chapter 5*, I consider the argument from moral disagreement, according to which the nature of moral disagreement should have us conclude that there are no moral properties. This argument can be proposed in support of either non-cognitivism or error-theory. Reductionism, I argue, fails to explain the relevant kind of moral disagreements. By contrast, realism is capable of doing so by employing the realist formula. Consequently, in contrast to reductionism, realism has the resources to refute the argument from moral disagreement and maintain the existence of moral properties.

In *chapter 6*, I consider the view that moral judgements involve reasons to perform actions. Rather than considering the much-debated question of whether moral judgements involve rational reasons, as rationalists claim, I focus on the more basic notion that moral judgements imply moral reasons. By considering fundamental normative questions, I argue that reductionism fails to account for this notion, whereas realism is able to do so by employing the realist formula. As a result, unlike

reductionism, realism is able to provide an accurate account of the normativity of moral judgements.

In *chapter 7*, I consider the internalist argument, according to which internalism concerning moral motivation together with the Humean theory of motivation entails that moral judgements do not consist in beliefs. This argument has been proposed in support of non-cognitivism. I argue that this argument can be refuted by rejecting internalism and adopting externalism. As a part of this reasoning, I consider Michael Smith's fetishist argument in favour of internalism and against externalism. Smith's argument raises the question of how people's moral motivation should be characterised according to externalism. I argue that realism, in contrast to reductionism, is able to account for how externalists should understand a pertinent aspect of moral motivation. Thus, similar to the reasoning in *chapter 5*, I maintain that realism, unlike reductionism, is able to account for a meta-ethical issue that is thought to provide support to non-cognitivism. In contrast to reductionism, realism has accordingly the resources to refute the internalist argument and maintain that moral judgements consist in beliefs.

Above I indicated that for realism to be a reasonable meta-ethical position, it is vital that it can account for our notion that objects have moral properties in virtue of having non-moral properties or, as I also will formulate it, that moral properties depend on non-moral properties. The importance of this is underlined by the fact that some of the arguments proposed in earlier chapters appeal to the realist formula. In *chapter 8*, I develop the realist notion of moral dependence further. Above all, I defend the realist formula against various objections. One important feature of the realist formula is that it entails strong supervenience. Simon Blackburn and Jaegwon Kim have suggested arguments to the effect that realists face dilemmas in regard to strong supervenience. I argue that there are reasons to believe realism can avoid these dilemmas. Another kind of contentions that partly concern the realist notion of dependence is J. L. Mackie's arguments from queerness in support of his view that there are no moral properties. I suggest that realism is not vulnerable to the queerness arguments and that the existence of moral properties therefore can be maintained. However, it is also pointed out that Kim's and Mackie's arguments might raise problems for realism that I cannot deal with in the present thesis.

In *chapter 9*, I consider a well-known argument proposed by Gilbert Harman, according to which moral properties do not figure in causal explanations, at least not if they are irreducible. He maintains that we therefore are not justified in believing in the existence of such properties. Against Harman's contention I argue that it is reasonable to assume that moral properties do figure in the relevant kind of explanations, even if they are irreducible, and that realists therefore are justified in maintaining their existence. I also argue that there is reason to believe that moral properties can have causal powers. However, in relation to the causal powers of moral properties realism might face difficulties that I am not able to discuss here.

## 6. *Concluding Remarks*

In this chapter, I have set the ground for the subsequent chapters by defending a conception of moral realism according to which this view says, in brief, that there exist moral properties that are irreducible to non-moral properties. Moreover, I have characterised the meta-ethical views implied by moral realism and the views that constitute their opposites. We have also found that two of these meta-ethical views—non-cognitivism and error-theory—have difficulties that justify us into considering whether realism or reductionism is more plausible.

In the last section, I mentioned some of the issues I discuss in subsequent chapters. However, I should also indicate some important issues that I do *not* discuss. First, although I consider the two forms of reductionism that I find most important, there are other forms I do not discuss, in particular so-called 'revisionist reductionism' and the view that has been labelled 'moral functionalism'.<sup>50</sup> Second, with the exception of some brief remarks in chapter 6, I do not consider rationalism. Third, and

---

<sup>50</sup> One well-known advocate of revisionist reductionism is Richard B. Brandt; see Brandt (1979), esp. chap. 1. Moral functionalism is proposed by Frank Jackson and Philip Pettit; see Jackson and Pettit (1995), pp. 20–40, Jackson (1998), chap. 5 and 6, and Jackson (2003), pp. 557–575. Moral functionalism may be a reductionist view, although it does not need to be (see Jackson and Pettit (1995), pp. 27–28, and Jackson (1998), pp. 141–142). However, Jackson and Pettit seem to be attracted to a reductionist version of functionalism. Moreover, moral functionalism seem to exist in both an analytic version and a synthetic version (see Jackson and Pettit (1995, p. 28). I think it can be maintained that the arguments directed against reductionism in the chapters to follow can be directed against a reductionist version of moral functionalism as well, but I will not argue for this claim here. For criticism of various aspects of moral functionalism, see e.g. Hatzimoysis (2002), pp. 10–22; Roojen (1996), pp. 77–81; Smith (1994), pp. 48–56, and Zangwill (2000), pp. 275–286.

most important, with the exception of some brief remarks in chapter 8, I do not consider epistemic issues. Fourth, there are important issues concerning the implications of realism, and which version of this position that is most reasonable, that I do not discuss. Moreover, there are important general philosophical notions that I make use of but leave by and large uncommented. Some of the issues I do discuss also require a good deal more attention than I give them. The main reason for this incompleteness is lack of time, space and ability. But I have also wanted to pursue a certain line of thought, something that might have been more difficult if I had considered additional issues. Since some of the issues I do not discuss might be relevant to the plausibility of realism, this view is in need of further examination than I provide in the chapters to follow.

## Chapter 2

# Analytic Reductionism and the Open Question Argument

### 1. *Introduction*

In the last chapter, I mentioned that reductionism exists in basically two forms: analytic reductionism and synthetic reductionism. For a long time, analytic reductionism was thought to be the only available kind of reductionism, and it is presumably still the one that is most established in meta-ethics. It therefore seems appropriate to start with discussing the plausibility of this form of reductionism. Moreover, the most renowned argument against this view is G. E. Moore's open question argument. The main purpose with the present chapter is to argue that an amended version of this argument provides reason to believe that analytic reductionism is erroneous.<sup>1</sup>

In the next section, I provide a bare outline of analytic reductionism. In section 3, I revise the original version of the open question argument in order to make it as plausible and resistant against objections as possible. In sections 4 and 5, I develop the argument further and defend it against the most prominent objections raised against it. The conclusion of my reasoning is that the open question argument offers the best explanation of our responses to the questions put in the argument, namely that analytic reductionism is mistaken.

### 2. *Analytic Reductionism*

Analytic reductionism states that a term, 'M', applying to objects with a moral property M, has the same meaning as a term, 'G', applying to objects with a non-moral property G. Because of this meaning equivalence, advocates of this view claim that 'M' and 'G' refer to the same property and that M and G consequently are identical. Analytic reductionism can be explicated in the following way. According to the received view, if two general terms have the same meaning, then both terms apply to objects with a certain property and only to objects with this property, and this holds

---

<sup>1</sup> This chapter corresponds, with the exception of some minor revisions, to Strandberg (2004), pp. 179–196.

merely in virtue of language. Analytical reductionists are then committed to the truth of the following biconditional:

It is analytically necessary that, for any object  $x$ ,  $x$  is  $M$  if and only if  $x$  is  $G$ .

The biconditional says that it holds analytically that those and only those objects that have  $M$  also have  $G$ ; that is, merely as a matter of language, the terms ‘ $M$ ’ and ‘ $G$ ’ apply to objects with these properties in such a way that something is  $M$  if and only if it is  $G$ . The expression ‘necessary’ in the biconditional indicates that, since analytic truths like this hold true merely in virtue of language, they hold true whatever other, non-linguistic, conditions may be the case, i.e. in whatever counterfactual circumstances or possible worlds. As we saw in the last chapter, if a necessary biconditional of this kind holds, a necessary condition for  $M$  and  $G$  being identical is satisfied. According to a somewhat simpler way of formulating analytic reductionism that I will occasionally make use of, it says that the meaning of a moral term ‘ $M$ ’ is constituted by a non-moral property  $G$  with which  $M$  is identical.

Analytic reductionists are generally considered to propose reductive analyses of moral terms, and I will at times employ that usage here. It might be argued, however, that analysis should be distinguished from sameness of meaning. That an analysis expresses sameness of meaning is arguably a necessary condition of its being correct, but this condition may not be sufficient to guarantee such correctness. Here I will be interested in analysis only in so far as it concerns sameness of meaning.

In the subsequent sections, there will be opportunities to further clarify what analytic reductionism amounts to.<sup>2</sup>

### 3. *An Amended Version of the Open Question Argument*

In *Principia Ethica*, Moore presents an argument that has been thought to demonstrate that analytic reductionism is erroneous: the open question argument. For a long time, and for several reasons, this argument has however become subject to several important objections. Partly as a

---

<sup>2</sup> Whether a certain work represents analytic reductionism is matter of interpretation, but among others the following ones are relevant: Adams (1981 (1973)), pp. 83–108; Firth (1952), pp. 317–345; Harman (1975), pp. 3–22; Lewis (1989), 113–137; Smith (1994), esp. chap. 5; Wong (1984), esp. chap. 4, and Zimmerman (2001), chap. 4. See also Jackson and Pettit (1995), pp. 20–40; Jackson (1998), chap. 6, and Jackson (2003), pp. 557–575. It should be noted, however, that some of these authors focus on values in general rather than on moral properties.



consequence of these objections, analytic reductionism has become more popular in recent years. However, I think the open question argument can be developed in a way that makes it possible to answer these objections and that the argument, thus understood, indicates the falsity of analytic reductionism.

Moore formulates the argument in the following well-known passage:

The hypothesis that disagreement about the meaning of good is disagreement with regard to the correct analysis of a given whole, may be most plainly seen to be incorrect by consideration of the fact that, whatever definition be offered, it may be always asked, with significance, of the complex so defined, whether it is itself good. To take, for instance, one of the more plausible, because one of the more complicated, of such proposed definitions, it may easily be thought, at first sight, that to be good may mean to be that which we desire to desire. Thus if we apply this definition to a particular instance and say 'When we think that A is good, we are thinking that A is one of the things which we desire to desire', our proposition may seem quite plausible. But, if we carry the investigation further, and ask ourselves 'Is it good to desire to desire A?' it is apparent, on a little reflection, that this question is itself as intelligible, as the original question 'Is A good?'—that we are, in fact, now asking for exactly the same information about the desire to desire A, for which we formerly asked with regard to A itself. But it is also apparent that the meaning of this second question cannot be correctly analysed into 'Is the desire to desire A one of the things which we desire to desire?': we have not before our minds anything so complicated as the question 'Do we desire to desire to desire to desire A?' Moreover any one can easily convince himself by inspection that the predicate of this proposition —'good' — is positively different from the notion of 'desiring to desire' which enters into its subject: 'That we should desire to desire A is good' is not merely equivalent to 'That A should be good is good.' It may indeed be true that what we desire to desire is always also good; perhaps, even the converse may be true: but it is very doubtful whether this is the case, and the mere fact that we understand very well what is meant by doubting it, shews clearly that we have two different notions before our minds.<sup>3</sup>

A reasonable understanding of Moore's argument is this. If an analysis is proposed to the effect that 'good' has the same meaning as a certain term or sequence of terms, e.g. 'desire to desire', one should ask 'Is whatever is desired to be desired good?' And if one finds the question intelligible—that is, if the question is 'open'—it follows that the question, and hence the analysis, does not express sameness of meaning. Thus, the analysis fails in this regard.<sup>4</sup> This is the *first part* of the argument. In the *second part* it is claimed that there can be no successful reductive analysis of 'good', since 'whatever definition be offered', we will find the corresponding question intelligible.

---

<sup>3</sup> Moore (1993 (1903)), pp. 67–68.

<sup>4</sup> There has been some controversy about how to interpret the quoted passage; see e.g. Fumerton (1990), pp. 68–69, and Lewy (1965), pp. 258–259. The open question argument is closely related to the naturalistic fallacy. On this notion, see Frankena (1976 (1939)), pp. 1–11.

The argument should, I think, be modified in several ways.

Firstly, Moore applies the open question argument to ‘good’, whereas I will, as has become customary, apply it to ‘thin’ moral property terms. I have particularly in mind ‘morally right’, ‘morally good’, and their counterparts ‘morally wrong’ and ‘morally bad’. The main reason is of course that I am concerned with whether moral properties referred to by these terms are reducible. Another reason is that it can be questioned whether it makes sense to say that something is right or good without presupposing that it is right or good in a particular way, e.g. morally, rationally or prudentially.

Secondly, we should modify the response to the question figuring in the argument. Moore suggests that the relevant response is that one finds the question intelligible (or ‘significant’, which I take to amount to the same thing). To find a question intelligible is presumably to take it to be raised with some point. However, I believe that this response is not entirely suited to the purpose of the argument, since it is possible to find a question intelligible in this sense even if it involves sameness of meaning. For instance, the question ‘Is whatever is a male that has a sibling who has a child an uncle?’ seems intelligible. As we shall see below when we discuss the paradox of analysis, such questions may be intelligible because, although they express analytic truths, they convey information in some sense.

To avoid this problem, I suggest that we take the relevant response to the question to be *doubt* as regards the correct answer—a response Moore hints at towards the end of the quoted passage. The basic idea is that if a truth is analytic we should not, as persons with a particular kind of knowledge, have doubt as to how to answer a question corresponding to it, even if we may find it intelligible. Although the question ‘Is whatever is a male that has a sibling who has a child an uncle?’ is intelligible, we do not doubt how to answer it, at least not after some reflection. I should also mention that I use the term ‘doubt’ in a somewhat unusual way, in that I take it to cover a spectrum of certainty in reactions to the question, ranging from doubt in the proper sense (where one is not sure whether to answer affirmatively or negatively), to the propensity to respond with an outright ‘no’. The reason for using ‘doubt’ in this broad sense is that if a response of hesitation—doubt proper—signals that the analysis is incorrect, so does a

clear ‘no’.<sup>5</sup>

Thirdly, it should also be pointed out that the persons whose doubt is alleged to be significant have to be linguistically competent with respect to the terms in question; that is, they have to be competent language users of this part of language. If they are not, it would be pointless to ask them a question with the purpose of investigating whether the terms have the same meaning. We should however not regard such a person’s doubt as an indisputable proof of the incorrectness of the proposed analysis. On occasion, even a person who is competent in the required way can be mistaken about what a term means, so we should say that the presence of doubt merely suggests that the analysis is incorrect.

Fourthly, we should increase the number of questions put to use in the argument. As we saw above, the argument invites us to take the non-moral property on the right-hand side of an allegedly analytically necessary biconditional, G above, and inquire whether it is a sufficient condition for something having the moral property on the left-hand side, M: ‘Is whatever is G also M?’ Since a successful analysis of the relevant kind requires the truth of a biconditional, it follows that if there is a response of doubt indicating that G is not a sufficient condition, the analysis fails. But there is no reason to confine the argument so that it employs only questions of this type. Given that the response to this type of question is significant, we could reach the same conclusion if we instead ask ‘Is whatever is M also G?’, i.e. ask if G is a necessary condition for something having M, and the response

---

<sup>5</sup> There seems to be a further reason to view the response of doubt as regards the correct answer to the question, not of finding the question intelligible, as the relevant response. Whether one finds a certain question intelligible or not presumably depends partly on what one believes about the person who asks it: what one believes about her knowledge, intelligence, sense of humour, etc. It would seem that a question may be found intelligible even if it contains an analytic truth, since it can be surmised that it is asked by a person who, for example, is not linguistically competent in the relevant way or who wants to make some strange kind of joke. Finding the question intelligible can then not be used as a test of whether it contains an analytic truth. Of course, one could stipulate that one always should imagine the question to be asked by a person who is linguistically competent, who does not want to make a joke, etc. However, the problem would then be that to employ this response is unnecessarily convoluted since it forces one to take a stand on things that are irrelevant to whether the question contains an analytic truth, namely, among other things, the epistemic position and intentions of the person who is imagined to be asking it. If, on the other hand, one uses doubt as the relevant response, one does not have to advance any particular hypotheses about the person who asks the question, for one simply asks *oneself* the question and lets *one’s own* doubt determine whether the question contains an analytic truth. To the extent that one is linguistically competent with respect to the terms in question, one’s response is reliable. We will then have a response that is better attuned to what the argument is intended to test.

is one of doubt. Similarly, where G is complex—so that it follows that if something has G it has e.g. F, H and I—we could ask ‘Is whatever is M also F?’ (i.e. ask if a certain ‘part’ of G is a necessary condition for something having M) and so on. Generally, we can take propositions which would be analytic truths if the proposed analysis is correct, reformulate them into questions, and allow our responses to these questions to indicate whether the analysis states sameness of meaning.

Lastly, we should be more careful than Moore concerning the second part of the argument. The first part of the argument has not been applied to all possible analyses of moral terms. Even if one accepts this part of the argument, it is therefore difficult to see how one can rightly claim to know that ‘whatever’ analysis is offered, one’s response would be one of doubt. One reaction to this challenge is to say that the argument only offers a test. A more offensive rejoinder would be to maintain that the claim that there is no correct reductive analysis offers the best explanation of the evidence available. We will come back to this issue in section 5.

We will have further opportunities to clarify this modified version of the open question argument when we consider objections to it. Before we do this, we should however consider whether there is any reason in the first place to think that the procedure of inquiring into the meanings of terms that the argument makes use of is reliable. In other words, if a linguistically competent person is asked a question in which a proposed analysis is contained—or what would be an analytic truth if the analysis is correct—and she has doubts about what to answer, does this suggest that the analysis fails to express sameness of meaning?

To answer this question, we may start by considering two biconditionals.

- (1) x is a brother if and only if x stands in the family relation to a person that is the main theme of the story of Cain and Abel.

While true, (1) does not state the meaning of ‘brother’. An indication of this is that we can imagine that ‘brother’ applies to a person even in counterfactual circumstances in which the right-hand side of (1) is false. By contrast, consider:

- (2) x is a brother if and only if x is a male sibling.

An indication of (2) being analytically necessary is that one cannot think of any object in any counterfactual circumstance such that ‘brother’ is, and

‘male sibling’ is not, attributable to it (or vice versa).

It might be argued that there are good reasons for us to use language in such a way as to include analytic truths like (2), although they may not be as strong as biconditionals. However, I would like to stress that I do not want to commit myself to any particular theory about analyticity. Rather, I here speculate about what one takes the function of analytic truths to be if one thinks there are such truths. It might be argued, however, that in order to communicate properly we have, at least regarding some terms, to agree that when the term is used, it picks out the same set of properties (or at least a limited set of a range of properties), in all counterfactual circumstances merely in virtue of language. First, if some terms did not pick out the same set of properties (or a limited set of a range of properties) in this way, it would be difficult to understand how communication was at all possible, since we would not have any guarantee that we were talking about the same things or the same properties of things. Moreover, it is important that truths about which properties these terms pick out are not conditioned by factual circumstances but hold merely in virtue of language. One reason is that if a term is to be useful in communication, it is important that persons with quite different knowledge of the world understand which properties the term picks out. If the correctness in applications of a term were conditional on the way the world actually has been, is, or will be in certain respects, the term could not be used by persons who lacked knowledge of these particular facts. Hence (1), unlike (2), would be rather useless as a rule for the usage of ‘brother’. Analytic truths, then, can be said to work as fixed points in language: they guarantee that we talk about the same things and the same properties of things in different circumstances, and that we do this from a common epistemological basis ensuring mutual understanding.

Now, how do these simple remarks bear upon the question whether the open question argument can be relied on in inquiring into the meanings of terms? If some truths hold merely in virtue of language, and if the existence of such truths is an essential feature of us being able to communicate, it seems reasonable to assume that persons who are linguistically competent possess knowledge of these truths. Indeed, if they did not, they would lack something essential in their ability to communicate, and it could then be questioned whether they really were competent in the relevant respect. But if linguistically competent persons

have this knowledge, it seems reasonable to treat their responses of doubt to questions of the kind put in the open question argument, such as ‘Is whatever is M also G?’, as indications of ‘M’ not having the same meaning as ‘G’. For these responses would simply be expressions of the linguistic competence of these persons. It is this tight connection—that is, between analyticity and linguistic competence, and between linguistic competence and responses to the questions utilised in the argument—that gives us reason to believe that the open question argument makes use of a reliable procedure of inquiring into the meanings of terms.

#### 4. *Defence of the First Part of the Argument*

(i) One well-known objection to Moore’s argument can be formulated as the following *reductio*. If the open question argument were valid, it could be used to refute hypotheses to the effect that two terms refer to the same property or natural kind. But it can clearly not do this.<sup>6</sup> To take Gilbert Harman’s example, suppose that we ask someone ‘Is whatever is H<sub>2</sub>O water?’ and that the person is doubtful of the correct answer. Her response would however not indicate that ‘water’ does not refer to the same natural kind as ‘H<sub>2</sub>O’.<sup>7</sup>

This objection rests on the presupposition that the open question argument is intended to determine whether two terms refer to the same entity, e.g. the same property or natural kind. Although this might have been how Moore understood the argument, we may, as I suggested in the last section, consider it to indicate whether two terms refer to the same property in virtue of having the same meaning (in the indicated sense of ‘meaning’).<sup>8</sup> The argument does not necessarily have any implications for the question whether two terms refer to the same entity where the reason for this is understood in some other way.

(ii) According to another objection to the open question argument, the fact that a person does not find a question intelligible—or does not have

---

<sup>6</sup> See e.g. Brink (1989), pp. 164–166; Durrant (1970), pp. 360–361; Harman (1977), p. 19; Lycan (1986), pp. 80–81; Putnam (1981), pp. 206–208, and Sturgeon (2003), pp. 533–534. For defences of Moore’s argument against this contention, see Baldwin (2003), pp. 322–324; Ball (1988), pp. 198–201, and Ball (1991), pp. 8–15. For another positive view of Moore’s argument, see Shafer-Landau (2003), pp. 56–58.

<sup>7</sup> Harman (1977), p. 19.

<sup>8</sup> Moore’s view of the argument was presumably due to his assumption that two terms refer to the same property only if they have the same meaning. For relevant references, see previous note.

doubts about the answer—does not show that the question expresses sameness of meaning.<sup>9</sup> For instance, the fact that we do not doubt how to answer the question ‘Is whatever is H<sub>2</sub>O water?’ does not show that ‘water’ and ‘H<sub>2</sub>O’ have the same meaning. Quite generally, there seem to be a number of truths we do not doubt but that are not analytic.

This objection rests on a misunderstanding. The open question argument does not imply that lack of doubt is a sufficient condition for sameness of meaning. Thus, the fact that we do not doubt how to answer the mentioned question does not imply that ‘water’ and ‘H<sub>2</sub>O’ have the same meaning. According to the first part of the open question argument, if we, in our capacity as linguistically competent, doubt how to answer a question in which a proposed analysis is contained, this indicates that the analysis fails to express sameness of meaning. Thus, the argument rather suggests that lack of doubt is a necessary condition for sameness of meaning. However, as we have already seen and as we will see further below, this claim should be qualified in various ways.

(iii) Some authors suggest that advocates of the open question argument overlook the paradox of analysis.<sup>10</sup> One way to formulate this paradox is as follows. A successful analysis should express sameness of meaning. In doing this, an analysis cannot state anything that is not contained in what is being analysed, since it should report only what holds merely in virtue of language. In principle, then,

(3) x is an uncle if and only if x is male and has a sibling who has a child,

should not state anything that is not contained in

(4) x is an uncle if and only if x is an uncle.

So, on the one hand, successful analyses should be uninformative, as (4) is. But on the other hand, it is obvious that, even when they are considered to be correct, they are often informative. To take the minimal case, (3) is at least somewhat informative.

The paradox of analysis is chiefly a problem for the pursuit of analysis as such, but it can also be appealed to against the open question argument in the following way. The most reasonable response to the paradox is generally

---

<sup>9</sup> See Brandt (1959), pp. 165–166. For a criticism of Brandt’s view, see Nielsen (1974), pp. 51–56.

<sup>10</sup> See e.g. Baldwin (1990), p. 88; Fumerton (1983), pp. 477–479; Fumerton (1990), p. 73, and Smith (1994), pp. 37–38.

considered to be that analyses can be correct and yet informative. However, the open question argument may seem difficult to reconcile with this view. In its original formulation, the argument suggests that we can take the fact that an analysis is such that, when it is reformulated into a question, the question is found to be intelligible as an indication of the incorrectness of the analysis. The argument may thereby be taken to imply that an analysis cannot be correct and yet informative. This is indeed so, since if an analysis is informative, we find a question corresponding to it intelligible. Quite generally, if a question conveys information that is new to us, we find it intelligible.

It should be noted, however, that this objection is not effective against the open question argument as it is formulated here, where the relevant response is that of *doubt* or lack thereof. For it is possible for an analysis to be correct and informative and yet such that, when it is reformulated into a question, one does not doubt how to answer the question. Understood in this way, the open question argument is thus compatible with the recommended response to the paradox of analysis. In general, there seems to be no inconsistency in claiming that one does not doubt the truth of a sentence even if the sentence is informative. However, it might be argued that even if straightforward inconsistency is not involved, the fact that something like it occurs needs to be explained. I will postpone this explanation until another, and more pivotal, issue has been addressed.<sup>11</sup>

We should also bear in mind here that in the amended version of the open question argument, a wider set of questions is allowed so that not only a whole analysis can figure in a question but also parts thereof. The implication of this in the present context is that even if an analysis is correct and informative, and one is doubtful about what to answer to a question in

---

<sup>11</sup> In relation to the paradox of analysis a possible misreading of the open question argument can be mentioned. (See Fumerton (1983), p. 478, and Magnell (1988), pp. 77–78.) In order to preserve sameness of meaning (3) above should not report anything that is not involved in (4). Now, the open question argument says that if one responds with doubt to an analysis in the interrogative form, this indicates that the analysis fails to express sameness of meaning. It may then be thought that the argument implies that if two sentences, such as (3) and (4), reformulated into questions, differ in terms of the epistemic attitude they elicit, as (3) and (4) probably would, then they have different meanings (which (3) and (4) do not). However, a defender of the argument is not committed to the view that the two questions have to elicit the same response in order to contain sameness of meaning. Even if the questions contain sameness of meaning, our responses to them may vary from firm certainty to other considerably weaker reactions as long as doubt, in the sense described above, is not involved.



which the whole analysis occurs, one may not be doubtful about what to answer to a question in which only a part of it occurs. What it is mentioned below about reflection is also relevant here. If one's response to a question put in the open question argument is allowed to be grounded in reflection upon it, a question expressing sameness of meaning may elicit a doubt-free response.

However, as already indicated, there is an issue regarding the open question and the paradox of analysis that needs to be addressed. The favoured response to the paradox of analysis is, as mentioned, that an analysis can be both correct and informative. But, as was argued in the last section, the open question argument receives support from the notion that the reason why a person's doubt how to answer a question put in the argument has anything to say about sameness of meaning is that she, *qua* linguistically competent, knows analytic truths. Now, it might be wondered how the response to the paradox squares with this aspect of the argument. If a person knows a certain truth, how can it be informative for her?

The answer to this question is that something that one, in one sense, knows can indeed be informative. In order to see this, we have to make a distinction between two kinds of knowledge of the meanings of terms: knowledge *how* and knowledge *that*.<sup>12</sup> When we say that a person knows the meaning of a term, we are normally referring to the fact that she is able to apply the term correctly. If the meaning of a certain term 'M' can be stated by a biconditional, we can say that she applies 'M' in accordance with the biconditional. This is an instance of knowledge *how*. It is knowledge of how to use a term, i.e. the ability to apply it correctly. This kind of knowledge should be separated from a more comprehensive kind of knowledge. A person endowed with the mentioned ability may start to reflect on it and try to formulate a principle that codifies the pattern she follows in using the term. That is, instead of merely applying the term according to the biconditional, she may try to formulate it explicitly. After such reflection, she may have gained propositional knowledge of an analytically true biconditional. She will now have knowledge *that* regarding

---

<sup>12</sup> This distinction is found explicit or implicit in writings about the paradox of analysis; see e.g. Fumerton (1983), pp. 488–489; King (1990), pp. 162–163, and Myers (1971), pp. 301–302. Frank Snare has appealed to it in a paper addressing the open question argument: Snare (1975), p. 126. The distinction was presumably introduced by Gilbert Ryle: Ryle (1966 (1949)), chap. II.

the meaning of the term.

The objection mentioned above can now be answered. An analysis may be informative for a person who has knowledge *how*, but not knowledge *that*, in regard to a term's meaning. Even if she can apply the term correctly, she will not have knowledge *that* regarding its meaning. Consequently, when an analysis is presented to her, it is informative for her. There is an illuminating analogy to this phenomenon in our knowledge of grammar.<sup>13</sup> A person who is said to know a certain grammatical rule is able to apply the rule correctly. Such a person can then be said to have knowledge *how* that reflects the grammatical characteristic described in the rule. Yet, if she is presented with a formulation of the rule, it may be informative for her because she lacks knowledge *that* regarding it.

The distinction between knowing *how* and knowing *that* could also be engaged in giving the explanation announced above. That is, it could explain how it is possible for a correct analysis to be informative and yet such that, when it is formulated as a question, one does not have doubt as to how to respond to the question. By invoking her knowledge *how* regarding the meaning of a certain term, a person may be in no doubt regarding the answer to the question, for the analysis accords with her use of the term. However, the analysis may yet be informative for her if she lacks knowledge *that* regarding the term's meaning.<sup>14</sup>

Having noticed the difference between the two kinds of knowledge in relation to meaning, we can also be clearer about what the open question argument amounts to. Questions that figure in the argument, such as 'Is whatever is M really G?', should be assumed to be addressed to persons who are linguistically competent in virtue of having knowledge *how* regarding the meanings of terms used in the questions. In the present case the addressee is asked whether things with the property in virtue of which 'M' applies, i.e. M, must have the non-moral property G, in virtue of which 'G' applies. The argument is devised to appeal to how such a person actually would apply 'M'. It aims to investigate whether, exercising knowledge *how*, she would apply it in this way. Hence, the argument does not require the

---

<sup>13</sup> Cf. Snare (1975), pp. 126–127.

<sup>14</sup> Hilary Putnam stresses that the open question argument should be understood to employ the ability to use terms, not abstract knowledge of concepts; see Putnam (1975a), p. 13.

addressee to possess knowledge *that* regarding the meaning of the term.<sup>15</sup>

That it is knowledge *how* as regards meaning that is involved in the open question argument also reveals that the procedure utilised in the argument is involved in everyday situations—situations in which one in a number of ways and for different reasons asks questions about the meanings of terms. Suppose, for example, that I am unsure whether I have used a certain term, e.g. ‘racist’, correctly. In order to check this, I may employ the procedure used in the open question argument and by doing so utilise my knowledge *how* regarding the meanings of terms. I may ask, for example, if a person who has the characteristics that prompted me to use the term really is a racist, i.e. if she has properties in virtue of which ‘racist’ applies. Or suppose that I am trying to find a word for something I want to express, but that I have difficulty doing so. In order to find a suitable term, I am likely to come up with various proposals. When I examine these proposals, I may employ the procedure used in the open question argument and by doing so utilise my knowledge *how* regarding the meanings of terms. For example, I may ask if something to which a proposed term, say ‘racist’, applies has the properties I want to pick out.

As these examples demonstrate, the kind of questions persons with knowledge *how* ask in order to inquire into the meanings of terms remains the same, and the kind of response, doubt or lack thereof, has the same function, irrespective of which the particular terms at issue are. Thus, the examples emphasise that the open question argument is just an instance of a general procedure of inquiring into the meanings of terms.

These examples also suggest that one can execute these inquiries with various degrees of reflection. In some cases—of which the second case above might be an example—one may perform the procedure routinely without being conscious of asking various questions and responding to them. In other cases, as when counterexamples are involved, these inquiries may be executed with a considerable degree of reflection.<sup>16</sup> In the latter

---

<sup>15</sup> However, after having doubted the kind of question that is asked in the open question argument, a person may acquire a kind of knowledge *that* regarding the meaning of a term: the knowledge that the term does not have the meaning proposed in the analysis at issue.

<sup>16</sup> The procedure used in the open question argument is intimately, although not necessarily, linked to counterexamples. In some cases discovery of a counterexample can be said to be a part of the argument. When one asks ‘Is whatever is M also G?’, it comes naturally to try to imagine some existing or non-existing object that is M but not G, and so on for other questions the argument employs. If one can do so, one has found a

kind of cases, one may be highly aware of the procedure, reflect critically on the different questions asked, their implications, and one's own responses. As a result of such an inquiry, one might at best come up with a correct analysis of the term. It is moreover reasonable to assume that the degree of reflection needed to reach a significant result may vary depending on the term at issue. There seems to be no reason not to acknowledge that there will also be various degrees of reflection when the issue concerns the meaning of a moral term. If the open question argument is an instance of a general way of inquiring into the meanings of terms that may vary as regards degree of reflection, the argument itself may take different forms in this regard.

The fact that the open question argument is part of a general procedure of inquiring into meaning has implications for the strength of the argument. If the procedure utilised in the argument is brought to bear on non-moral cases, and if it is significant in these cases, then someone who wishes to argue that the open question argument does not have any implications for the meanings of moral terms will need to point to a difference between the moral case and other cases. (As we shall see below, there is at least one objection along these lines.) Of course, an objector to the argument may maintain that the procedure has no implications for meaning at all. That, however, would be a very strong claim, blatantly at variance with the fact that we do use the procedure and in fact rely on it to a great extent.

(iv) Some authors accept that it is possible to inquire into the meanings of certain terms by means of the kind of questions put in the open question argument. They argue however that in certain cases, of which the moral case is one, a person may find such a question intelligible—or may have doubts about the answer—even if the analysis is correct. The reason they submit is that an analysis may be so complex that the resulting question is difficult to grasp fully, which explains why one's response to it does not

---

counterexample to the proposed analysis, and as a result, one responds to the question with doubt. In other cases a counterexample is not part of the argument, but it can be said to be tracked by the response. After having responded with doubt to a relevant question, one can try to find a counterexample on the assumption that one's response suggests the direction in which a counterexample is to be found. In yet other instances of the argument counterexamples may be absent.

accord with its involving an analytic truth.<sup>17</sup> Take the question:

- (5) Is whatever is a closed plane figure, equidistant at all points from some fixed point, a circle?<sup>18</sup>

It can be argued that although the analysis (5) contains is correct, one would have doubts about the answer because the analysis is so complex that (5) is difficult to grasp completely.

As a first reply to this objection, it should be recalled that in the open question argument as we now understand it, room is left for reflection before one responds to the question. It is not the case, then, as this objection suggests, that one has to respond instantaneously without any deliberation. Thus, after reflection one may come up with a response to (5) that is significant as to the correctness of the proposed analysis.

It should further be noted that this objection presupposes that the open question argument only can employ questions of one type: questions in which it is asked if the right-hand side of the analysis, *analysans*, is sufficient for the left-hand side, *analysandum*. Above we extended the range of possible questions, however, to include questions about whether the whole or part of *analysans* is necessary for *analysandum*. Since it is not a complete analysis that figures in the latter type of questions, such a question is presumably easier to grasp; and given this, a response of doubt to such a question can be taken to indicate that the analysis fails to express sameness of meaning. Let us grant that (5) is difficult to grasp owing to its complexity. The questions ‘Is whatever is a circle equidistant at all points from some fixed point?’ and ‘Is whatever is a circle a plane figure?’ may be easier to grasp and not elicit doubt.

There is a further point that is relevant to the plausibility of the objection under consideration. As was noted above, the procedure of inquiring into the meanings of terms employed in the open question argument is also utilised when an analysis is devised. Proponents of the present objection claim that an analysis might be so complex that the procedure employed in the open question argument is not applicable in the relevant way. However, they have then difficulties explaining how a correct complex analysis can be formulated at all.

---

<sup>17</sup> See Brandt (1959), pp. 164–165; Broad (1971 (1934)), p. 115; Jackson (1998), p. 151, and Zimmerman (2001), pp. 78–79.

<sup>18</sup> I have borrowed this example from Zimmerman (2001), p. 79.

### 5. *Defence of the Second Part of the Argument*

If what has been said so far is correct, it follows that when an analysis of a moral term is proposed, one can put the analysis—or what should be an analytic truth if the analysis is correct—in the interrogative form, and if one, as a person who is linguistically competent with respect to the terms at issue, has doubt (in the described sense and perhaps after reflection) about how to answer the question, this suggests that the analysis fails to express sameness of meaning. The open question argument can in other words be used as a test of whether an analysis states sameness of meaning. This corresponds to how Moore uses the argument in chapters II, III and IV of *Principia Ethica*. However, as is indicated by the quotation above, Moore seems to have thought that the argument offers not just a test for individual analyses but also generally shows that analytic reductionism is mistaken. The latter is obviously a much stronger claim and there is reason to take up a humble attitude towards efforts to prove it outright. But although no such straightforward proof may be available, it is plausible to argue that the claim that there is no correct reductive analysis of moral terms offers a better explanation of available evidence than does the contrary claim.

Let us first recall that the amended version of the open question argument has extended the range of responses that can indicate whether an analysis is correct or not. Besides asking whether the analysis offers a sufficient condition, we can ask if it, or a part of it, offers a necessary condition. Furthermore, besides taking into account instantaneous responses to the questions asked, we have allowed responses that to various degrees are based on reflection and its results, such as counterexamples. Given the wide range of questions and accompanying responses that result from these amendments, it is not far-fetched to hypothesise that, in meta-ethical debate to date, no analysis of a moral term has been presented that has not been exposed to a question of the relevant sort and has not also elicited a response of doubt to the question. Most of the presented analyses, especially the more prominent and discussed ones, seem in fact to have been the subject of a number of such questions, questions that have elicited responses of doubt. Moreover, in these analyses there seems to be no significant necessary condition that has not figured in a question the answer of which has been

doubted.<sup>19</sup>

Suppose we accept these generalisations. (Admittedly, they are somewhat bold but, as I have indicated, I think they can be defended in the light of the reception of various analyses.<sup>20</sup>) What we should ask then is: which is the best explanation of the available evidence? Two explanations are relevant here. The first is that there is a correct reductive analysis of the moral term we are interested in. The other is that there is no correct reductive analysis of the term in question.

Suppose first, in accordance with the first explanation, that there is a correct reductive analysis, but that it has not yet been presented. For this claim there seems to be no evidence. Until an analysis has been presented that, were it put interrogatively, we would not doubt, there is no reason to believe that there is a correct analysis of this kind.

Furthermore, advocates of this version of the first explanation have to explain why, in spite of the amount of the time and effort spent on the enterprise, no analysis has yet been presented which is not vulnerable to a response of doubt. In particular, they have to explain why no necessary condition has yet been presented which has not elicited this response. (Note that it would be misleading to point out here that there are few accepted analyses of non-moral terms even when it comes to rather trivial terms, such as ‘chair’, ‘ball’, etc. Any estimation of the prospect of accomplishing a correct analysis of a certain term will have to be related to the time and effort invested in the project.) Assuming that the philosophers who have proposed analyses are linguistically competent, in the sense of having knowledge *how* regarding the meanings of moral terms, it seems reasonable to expect that, however difficult the task might be, at least one necessary condition should have been presented by now that is not vulnerable to a sceptical response. It should also be noted that even if a necessary analytic

---

<sup>19</sup> The qualification ‘significant’ is important, although admittedly vague. There are analytically necessary conditions that have not been doubted, e.g. ‘If something has a moral property it has it in virtue of some of its non-moral properties’. This and other truths are often implied by proposed analyses, but they are not such that they together with other necessary conditions that have not been doubted can constitute correct reductive analyses. When I use the expression ‘necessary condition’ in the present context, I refer to a necessary condition that is significant in a way that these conditions are not.

<sup>20</sup> They are also supported by some more or less systematic surveys; see e.g. Feldman (1978), chap. 2; Pigden (1991), pp. 428–430, and Schurz (1997), chap. 11. Cf. Railton (1989), p. 152. Ball points out that the open question argument should be understood as ‘an *inductive* generalization from linguistic behaviour’ (Ball (1991), p. 19).

truth of the relevant kind is found, this would not necessarily take the analytic reductionist very far, since an analytic reduction would presumably consist of a number of such truths.<sup>21</sup>

Suppose, next, in accordance with the first explanation, that a correct analysis *has* been presented. This claim seems difficult to reconcile with the evidence of responses of doubt to questions containing analyses or parts of them. If the previous arguments are correct, then, if a question contains an analytic truth, one should not, given the qualifications entered, respond to it with doubt. But so far there seems to be no analysis that has not been responded to in this way, and as regards more prominent analyses, there seem to be a number of such responses. Of course, it might be replied that the open question argument does not support these claims because it does not indicate whether an analysis is correct. But I have argued, against some notable objections, that it does. Particularly, it is difficult to deny that it does given the connection between analytic truths, linguistic competence and responses to the questions put in the argument. As linguistically competent persons we have knowledge of these truths, and our responses to the relevant questions could then be taken to suggest that the proposed analysis is incorrect. Further, it is difficult to deny that the open question argument indicates whether an analysis is correct given that the procedure it utilises is prevalent in other, non-moral, cases. Denying the significance of the argument, one is committed to the view that an established procedure of testing meaning is generally mistaken. However, that contention would be hard to render plausible.

It should also be noted that advocates of this version of the first explanation face an explanatory task corresponding to the one mentioned above. Considering the invested time and effort, they have to explain why, given that there is a correct analysis, analyses to such an extent have been developed that are vulnerable to sceptical responses, even regarding their necessary conditions.

The second explanation avoids the problems of both versions of the first explanation. The hypothesis that there is no correct reductive analysis is supported by evidence—that there is no presented analysis that, when put interrogatively, has not been doubted, even as regards a necessary condition.

---

<sup>21</sup> In chapter 7, I will argue that one of the strongest candidates for a necessary condition does not constitute such a condition. Accordingly, this argument provides further reason to believe that there is no reductive analysis of the required kind.



And so the second explanation accounts for the fact that no such analysis has been presented: there is none. Besides its evidential and explanatory advantage, it is also simpler, since it does not require us to explain why no analysis has been presented that is invulnerable to a sceptical response, or—on the assumption that such an analysis has been presented—why so many analyses have been produced that are so vulnerable, even regarding their necessary conditions. There are thus reasons to conclude that the second explanation is to be preferred.

A certain objection against the second explanation may readily come to mind. It might be argued that also within other areas in philosophy there are no proposed analyses that have not been subject to the kind of doubt appealed to in the open question argument. But in these areas we normally do not take the best explanation to be that there is no correct reductive analysis. So why should we do so in meta-ethics?

It might actually be the case that we, after some consideration, would come to the conclusion that an explanation similar to the second one above is also correct for other areas. Nevertheless, we are not committed to accepting that kind of explanation in other cases only because we think it holds in the moral case. In fact, I think comparisons with other areas would confirm the view that the second type of explanation is especially appropriate in meta-ethics.

It should first be noted that theoretical submissions labelled ‘analysis’ in philosophy are often something else than what has been discussed above, i.e. an analysis of a term as it is used in ordinary language. For example, analyses might be intended to be revisionist. The fact that an analysis of the first kind has been shown to be unsuccessful is not necessarily a reason to think that there is no correct analysis of the latter kind.

Moreover, in other philosophical areas it seems easier to find at least some non-trivial necessary conditions that have not been seriously doubted. And even if these conditions have been doubted, this seems in any case less common than in meta-ethics. This suggests that, in contrast to meta-ethics, successful reductive analyses might be achieved in these areas. For example, I take it that it is fairly uncontroversial to claim that one acts freely only if one is not forced to do what one does not want to do, or is prevented from doing what one wants to do, by some internal or external power; and that one knows something only if one believes it, it is true and one is justified in

believing it. What is controversial is exactly how these conditions should be understood and whether they are sufficient. In meta-ethics, however, there seems to be no nontrivial necessary condition that has not been doubted, and that has not been doubted to a great extent.

Relatedly, in meta-ethics analyses are sometimes deliberately devised to be incompatible with features that are part of other analyses, features that equally deliberately have been incorporated in these analyses. For example, some meta-ethicists propose analyses designed to capture the view that the fact that an action is right or ought to be done gives a person reason to perform the action only if the person has a pro-attitude towards it. Others argue that this would imply that moral reasons are arbitrary and propose accordingly analyses designed to exclude precisely this feature.<sup>22</sup> Again, some meta-ethicists formulate analyses designed to capture the idea that something has a moral property only relative to a certain culture, society or group, whereas others propose analyses avowedly intended to exclude just this.<sup>23</sup> Such deliberate clashes between analyses do not seem prevalent in other areas. This suggests that there is little of conceptual common ground in meta-ethics and hence that the prospects of successful reductive analyses in the moral area are slim.

It is further notable that in other philosophical areas debate is often centred around counterexamples that move the debate forward. To describe the matter in a somewhat idealised way: an analysis of a certain term is proposed and a counterexample suggested; on the basis of the counterexample, the analysis is adjusted or a completely new analysis is proposed, and so on. The fact that the debate develops in this way suggests that there are quite specific conditions that have to be fulfilled for the terms to apply; otherwise the confidence in counterexamples would be difficult to explain. Although counterexamples are also common in meta-ethics, they tend there to be more controversial. They also tend to play a primarily negative role in that they are used to showing that a certain analysis is inaccurate, but do not move the debate forward in the way described. This suggests that as regards moral terms there are no particularly specific analytically necessary conditions that have to be fulfilled for the terms to apply and hence no conditions of the kind needed for there to be successful

---

<sup>22</sup> For example, compare Harman (1975), pp. 3–11, and Smith (1994), pp. 164–175.

<sup>23</sup> For example, compare Wong (1984), esp. chap. 4, 5 and 6, and Firth (1952), pp. 318–319.

reductive analyses of these terms.

## 6. *Concluding Remarks*

According to analytic reductionism, a term 'M', applying to objects with a moral property M, and a term 'G', applying to objects with a non-moral property G, have the same meaning and that M and G are identical. This view commits analytic reductionists to the truth of an analytically necessary biconditional which would establish that a necessary condition for M and G being identical is fulfilled. Above I have argued that the procedure employed in an amended version of Moore's open question argument can be used to test whether terms have the same meaning. Moreover, I have argued that the open question argument, thus understood, offers an explanation of our responses to the questions put in the argument that seems superior to its competitor: that there is no reductive analysis that succeeds to state sameness of meaning. Hence, the open question argument provides reason to believe that the mentioned biconditional is false and that the alleged identity does not hold. To put the point in a slightly different way: there is reason to believe that the meaning of a moral term 'M' is not constituted by a particular non-moral property G. Consequently, there are reasons to believe that analytic reductionism is mistaken.

It is important to stress, however, that this conclusion might be falsified in the light of further considerations. Although the mentioned explanation appears to be the best of the available evidence, there might turn out to be correct reductive analyses. In subsequent chapters, there will be reasons to return to analytic reductionism and direct further criticism towards it. However, thus far the conclusion of the open question argument seems sufficiently established to encourage us to try to find an alternative to analytic reductionism. Traditionally the open question argument has been thought to provide support to non-cognitivism.<sup>24</sup> But, as we saw in the last chapter, this view is vulnerable to several objections that at least *prima facie* seem serious. In that chapter we saw that also error-theory appears to have serious shortcomings. For these reasons, we should investigate whether there is some other version for reductionism, or a kind of realism, that is preferable to non-cognitivism and error-theory.

---

<sup>24</sup> See e.g. Ayer (1952), chap. 6, and Hare (1952), chap. 5. However, it might be argued that non-cognitivism itself is vulnerable to the open question argument; see Miller (2003), pp. 47–51, and Smith (2001b), pp. 102–107.

## Chapter 3

### Synthetic Reductionism and the Two Questions

#### 1. *Introduction*

According to analytic reductionism, a term applying to objects with a moral property and a term applying to objects with a non-moral property have the same meaning. On the basis of such a meaning equivalence, advocates of this view claim that the moral property and the non-moral property are identical. In the last chapter, I argued that an amended version of the open question argument provides reason to believe that the relevant terms do not have the same meaning and that the alleged identity consequently does not hold. Following Saul Kripke's and Hilary Putnam's account of the reference of natural kind terms, it is now widely accepted that two terms can refer to the same property despite not having the same meaning. Hence, property identity can be the case even though it is not expressed in meaning equivalence. Advocates of synthetic reductionism employ this account to maintain that a moral property is identical with a non-moral property. As the open question argument is designed to show that property identity does not hold because the relevant terms do not have the same meaning, it cannot be used to challenge this kind of reductionism. In this chapter, I will argue that synthetic reductionism is mistaken. However, in the next chapter I will suggest that it also contains many ideas worth retaining by realism.

In the next section, I briefly outline the aspects of Kripke's and Putnam's account of the reference of natural kind terms that I think are relevant to synthetic reductionism. In section **3**, I consider how synthetic reductionism might be understood. In sections **4** and **5**, I examine synthetic reductionism by posing two questions that appeal to thought experiments similar to those Kripke and Putnam employ when they argue for their view. Our considerations in relation to these questions suggest that synthetic reductionism is mistaken. However, they also point towards a realist account of the way in which non-moral properties make objects have moral properties which will be developed in chapter **4** and onwards.

## 2. *Natural Kind Terms*

According to Kripke's and Putnam's account of the reference of natural kind terms, we are able to apply such a term correctly because there are certain characteristics that we associate with it and that typically are had by the stuff to which the term applies.<sup>1</sup> Characteristics of this sort enable us to pick out the natural kind in such a manner that we can apply the term correctly even if we do not have knowledge of the correct scientific account of the natural kind. In some respects, these characteristics function in a similar way to analytically necessary conditions: they help us to talk about things and to do so from a shared epistemic perspective that facilitates mutual understanding. To know that the term applies to what has these characteristics may even be considered to be part of what it is to be linguistically competent with respect to the term. Thus far there appears to be significant similarities between natural kind terms and the type of terms upon which analytic reductionism models its reductionist endeavour.

However, on this view, the mentioned characteristics do not constitute the meaning of a natural kind term. Thus, these characteristics do not yield a necessary biconditional such that there is reason to believe that what the term refers to is identical with a set involving these characteristics. What *does* yield a necessary biconditional of this kind is a scientifically identified stuff, a natural kind, which the mentioned characteristics help us to pick out. These characteristics fix the reference of the term and may consequently be called 'reference-fixing characteristics'. They can do so because the natural kind is causally responsible for these characteristics and so causally regulates our use of the term by means of them. The mentioned necessary biconditional does not state sameness in meaning. It is therefore not true merely in virtue of language and is consequently synthetically, not analytically, necessary. Accordingly, it is sometimes said that a synthetic identity has been found.

Kripke and Putnam argue for their view of the reference of natural kind terms by help of thought experiments that employ our linguistic intuitions concerning the application of such a term in different possible worlds.<sup>2</sup> Suppose there is a stuff that causally regulates our use of a natural

---

<sup>1</sup> Kripke (1980), esp. Lecture III, and Putnam (1975b), pp. 215–271. There are differences between Kripke's and Putnam's views of the reference of natural kind terms, but as these are not relevant to the present investigation, I will ignore them.

<sup>2</sup> Cf. Horgan and Timmons (1992b), pp. 161–162.

kind term in the way just outlined. Suppose further that science has found what this stuff consists in and that it fulfils the appropriate requirements on natural kinds. On the present view, the term refers to that stuff in compliance with a necessary biconditional and hence in all possible worlds. In our capacity as linguistically competent with respect to 'water', we should then be prepared to apply the term to that stuff, and only to that stuff, in all possible worlds. Especially, we should be so disposed irrespective of whether the mentioned characteristics are exemplified or not.

Let us employ thought experiments in relation to what has become the main example of a natural kind term: 'water'.<sup>3</sup> The stuff which we apply this term to typically has certain characteristics, such as being colourless, transparent, odourless, wet and liquid. We associate these characteristics with the term and apply it to the stuff that has them. That stuff is causally responsible for these characteristics and so causally regulates our use of the term. Science has found that this stuff has the chemical formula  $H_2O$  and that it fulfils the appropriate requirements on natural kinds. Suppose that we are linguistically competent with respect to 'water' and that we know this result from science. In order to examine whether 'water' refers to  $H_2O$  in the way Kripke and Putnam maintain, we may imagine two kinds of possible worlds and ask two corresponding questions. Imagine first a possible world where there is a stuff which has quite another chemical formula than  $H_2O$ : XYZ. However, this stuff has the characteristics mentioned above: colourlessness, etc. The inhabitants of that possible world have a certain term, 'water', with which they associate these characteristics, and they apply the term to what has them. The stuff that causally regulates their use of the term via these characteristics is XYZ. We can now ask the *first question*: if there is something that is XYZ in that world, is it water? We seem to have the intuition to answer 'no'. We would say that it certainly appears like water, but that it is not water. Hence, there is reason to believe that  $H_2O$  is a necessary condition for something to be water in that possible world (and that the mentioned characteristics do not constitute a sufficient condition). Moreover, we seem to have the intuition that we and the inhabitants of that possible world refer to different stuffs with our respective terms; whereas we refer to  $H_2O$ , they refer to XYZ. Let us next imagine a possible world where there is a stuff that is  $H_2O$ . However, in that world  $H_2O$  does not

---

<sup>3</sup>See Kripke (1980), pp. 128–129, and Putnam (1975b), pp. 223–226, 229–235.

have the mentioned characteristics: colourlessness, etc. We may also assume that the inhabitants of that world have a term, 'water', with which they associate the mentioned characteristics in the way outlined above, but that their use of this term is not causally regulated by H<sub>2</sub>O. We can now ask the *second question*: if something is H<sub>2</sub>O in that world, is it water? Here our intuition seems to be to answer 'yes'. We would say that it certainly does not appear like water, but that it is water. Hence, there is reason to believe that H<sub>2</sub>O is a sufficient condition for something to be water in that possible world (and that the mentioned characteristics do not constitute a necessary condition).

The support our responses to the two questions provides to Kripke's and Putnam's view of the reference of natural kind terms can be described in different ways. It seems however correct to claim that, on the assumption that our responses to the two questions are reliable, they provide reasons to believe that something is water if and only if it is H<sub>2</sub>O. Since our responses concern what is the case in other possible worlds, they moreover provide reasons to believe that this biconditional holds necessarily. On this ground, it may be maintained that 'water' refers to H<sub>2</sub>O and that water is identical with H<sub>2</sub>O. One way to put this view is to say that 'water' is a rigid designator in that it refers to H<sub>2</sub>O in all possible worlds. However, 'water' and 'H<sub>2</sub>O' do not have the same meaning. Consequently, the mentioned biconditional does not hold merely in virtue of language and is not analytically necessary. As I indicated above, it is reasonable to assume that it is part of being linguistically competent with respect to 'water' to know that the term refers to *some* given stuff in the indicated way, given that it fulfils certain appropriate requirements. Otherwise it might be difficult to understand how our responses to the questions posed in the thought experiments could be assumed to generate reliable results. But it is not part of this competence to know that the term refers to H<sub>2</sub>O. Rather, this is something we come to know from a scientific and empirical investigation. There is therefore reason to claim that the mentioned biconditional is synthetically necessary.<sup>4</sup>

---

<sup>4</sup> Sometimes this is expressed by saying that it is metaphysically necessary that water is H<sub>2</sub>O.

### 3. *Synthetic Reductionism*

Synthetic reductionism states that a term ‘M’, applying to objects with a moral property M, and a term ‘G’, applying to objects with a non-moral property G, refer to the same property and that M and G consequently are identical. However, unlike analytic reductionists, advocates of this view do not believe that ‘M’ and ‘G’ have the same meaning. Accordingly, they do not assert the truth of an analytically necessary biconditional involving M and G. Yet, they assert the truth of the following biconditional:

It is synthetically necessary that, for any object x, x is M if and only if x is G.

If a biconditional of this kind holds, a necessary condition for M and G being identical is fulfilled.<sup>5</sup>

Although the inspiration to adopt synthetic reductionism comes from the account of the reference of natural kind terms outlined above, advocates of synthetic reductionism are presumably not prepared to map over this view of reference to the moral domain without amendments. Before I examine synthetic reductionism, I will therefore consider on what points advocates of this view should modify it in order that it be appropriate to moral terms. In connection with these points, I will take the opportunity of clarifying this kind of reductionism. This is important also for realism, as I will argue in the next chapter that this view may adopt a number of the ideas in synthetic reductionism. However, my aim is not to provide a full account of this view and many simplifications are unfortunately necessary.

The first point concerns the role of analytically necessary conditions on applications of natural kind terms and moral terms, respectively. The reference-fixing characteristics associated with natural kind terms function, as we saw above, in some respects *as if* they were such conditions, although

---

<sup>5</sup> Several authors have argued that a moral property may be identical with a non-moral or natural property even though this is not expressed in meaning equivalence; see e.g. Brink (1989), pp. 151–167; Brink (2001), pp. 157–163; Durrant (1970), pp. 360–361; Harman (1977), pp. 19–20; Lycan (1986), pp. 80–81; Putnam (1975a), p. 280, and Putnam (1981), pp. 205–208. (Although Brink is inclined to the view that moral properties are irreducible to, but constituted by, natural properties, he holds the possibility of reduction and property identity open.) For other accounts, see Adams (1981 (1979)), pp. 109–119; Copp (2000), pp. 120–134; Merli (2002), pp. 214–231, Railton (1989), pp. 155–161; Railton (1993b), pp. 315–328, and Sayre-McCord (1997a), pp. 267–292. See also Scanlon (1998), pp. 11–13. Richard N. Boyd offers an account of the reference of moral terms akin to Kripke’s and Putnam’s account of the reference of natural kind terms: Boyd (1988), pp. 194–199, 209–212. The discussion of moral relativism in Miller (1985), pp. 507–556, is inspired by Putnam’s view of reference.



they are not. But at least as regards moral terms there seem to be some characteristics that constitute genuine analytically necessary conditions for such terms to apply.<sup>6</sup> For example, in relation to ‘right’, I think it is fairly uncontroversial that the following two claims provide such conditions: ‘If something has a moral property, such as rightness, it has this property in virtue of having certain non-moral properties’, and ‘If an action is right, there is a moral reason to perform it’. (I will return to these conditions in subsequent chapters.) A synthetic reductionist may also want to include other conditions in this category. When considering thought experiments in the next two sections, we should presume that the relevant analytically necessary conditions are satisfied.

However, in consideration of the open question argument, there is reason to believe that the analytically necessary conditions are insufficient for us to be able to apply a moral term correctly.<sup>7</sup> It may therefore be maintained that there are some characteristics that help us to do so which work as reference-fixing characteristics, analogous to how such characteristics work in relation to natural kind terms.<sup>8</sup> For example, it might be assumed that there are certain characteristics that we associate with the term ‘right’ and which typically are had by the actions to which we apply this term. These characteristics, we may suppose, enable us to apply the term correctly even if we are ignorant of the accurate account of the property. Some of these characteristics may also be such that knowing that ‘right’ applies to actions that have them is part of what it means to be linguistically competent with respect to the term. However, in contrast to analytically necessary conditions, they are not necessarily had by what is right. The reason why the mentioned characteristics can have this function is that actions that are right have a non-moral property which is causally responsible for these characteristics and so causally regulates our use of ‘right’ by means of them.

---

<sup>6</sup> Cf. Adams (1981 (1979)), pp. 112–113. Michael Smith believes that synthetic reductionists have to deny that there are any analytically necessary conditions for ‘right’ to apply, or, if synthetic reductionists accept that there are such conditions, that they ‘simply fail to take into account’ these conditions (Smith (1994), p. 32). I cannot see that Smith offers any arguments for these claims.

<sup>7</sup> Here and elsewhere, I will assume that at least sometimes we are able to apply moral terms correctly.

<sup>8</sup> Cf. Adams (1981 (1979)), pp. 113–114; Horgan and Timmons (1992b), pp. 163–164; Miller (2003), pp. 163–164; Railton (1989), pp. 162–163, and Sayre-McCord (1997a), p. 276.

Since the mentioned characteristics are not necessarily had by the objects to which a moral term applies, they need not be had by all such objects. Some of these characteristics are presumably had by most, although perhaps not by all of them. But other such characteristics might be quite particular and had merely by a smaller group of the objects to which the moral term applies. Concerning 'right', many of the characteristics presumably consist in various features pertaining to people's beliefs, attitudes and responses in relation to actions; for example: 'People approve of right actions being performed', 'People believe that it is of great importance that actions which are right are performed', and 'People encourage each other to do what is right'.

We can now see, in bare outline, how synthetic reductionism identifies a moral property with a non-moral property. Suppose, for example, that our use of 'right' is causally regulated by a non-moral property G by means of characteristics of the kind mentioned above. On synthetic reductionism, 'right' refers to G, and rightness is identical with this non-moral property. As a consequence, synthetic reductionists are committed to the truth of a necessary biconditional of the kind mentioned above. This biconditional is synthetically, not analytically, necessary. It is not part of being linguistically competent with respect to 'right' to know that the term refers to G, although it presumably is part of this competence to know that 'right' refers to *some* given non-moral property in the indicated way. One important consequence of this conception is that a linguistically competent person may refer to G with 'right' even if she is not aware of this fact and even if she believes that rightness consists in something quite else.

However, it should be stressed that synthetic reductionists presumably are prepared to claim all this only on condition that certain appropriate requirements as regards G are fulfilled. Below I will make some remarks with bearing on these requirements, but more needs to be said about them to provide a satisfactory account of synthetic reductionism.

A second point concerns the epistemic methods by which it is found out that a natural kind term refers to a certain natural kind and a moral term refers to a certain non-moral property, respectively. As regards natural kind terms, such accounts are arrived at via scientific and empirical methods. In regard to the moral case, it might be argued that such methods are not

available or at least play a more subordinate role. However, in the meta-ethical literature, epistemic methods have been suggested that are not necessarily scientific or empirical, although some of them might be. The most influential is coherentism, according to which a belief is justified in so far as it is part of a coherent set of beliefs.<sup>9</sup>

In relation to natural kind terms and moral terms as understood by synthetic reductionism, one aspect of justification is the ability to explain reference-fixing characteristics. Thus, in order for the view that ‘water’ refers to H<sub>2</sub>O to be justified, it should be possible to explain how this stuff is causally responsible for the reference-fixing characteristics we associate with the term and so how it can causally regulate our use of it. An analogous demand holds for the view that moral terms refer in the way synthetic reductionists maintain.<sup>10</sup> On the view that a moral term refers to a certain non-moral property, it should be possible to explain how this property is causally responsible for the reference-fixing characteristics we associate with the term and so how the property can causally regulate our use of it.<sup>11</sup> For example, if it is claimed that ‘right’ refers to a certain non-moral property G, it should be possible to explain how G is causally responsible for characteristics such as ‘People approve of right actions being performed’. In this way, synthetic reductionism may be able to explain some of the functions of ‘right’ and many of our everyday views associated with the term.

As we have seen, according to synthetic reductionism a moral term such as ‘right’ refers to a non-moral property G which causally regulates our use of the term via reference-fixing characteristics. However, this should not be taken to imply that all people who refer to G with ‘right’ have their use of the term causally regulated by G in the sense that they stand in immediate causal contact with G. Some people’s use of the term might be directly causally regulated by G in this way, whereas other people might apply it to roughly the same objects as the first group of people, although they are not directly causally involved with G. Accordingly, there need not be causally

---

<sup>9</sup> See e.g. Boyd (1988), pp. 199–202, 206–209; Brink (1989), chap. 5; Goldman (1988), chap. 5, and Sayre-McCord (1996), pp. 137–179. For a useful overview of various epistemological positions in ethics, see Tännsjö (1990), chap. 2.

<sup>10</sup> Cf. Railton (1989), p. 161–163.

<sup>11</sup> The point here is not that such explanations actually have to be provided, but that there have to be explanations of this kind, even if they as a matter of fact are not available.

explanations of all people's use of 'right' in terms of their being in immediate causal contact with G. Indeed, it might be the case that quite a few people stand in such a contact with G. However, also in the latter case people's use of the term is causally regulated by G, since their use has its origin in other people's causal interaction with G.

It should be pointed out that I have merely given a very rough and simple sketch of how the causal regulation of our use of moral terms might be viewed on synthetic reductionism. I have not commented on certain important aspects of this notion which a satisfactory theory must account for.<sup>12</sup> For example, a theory of causal regulation should account for the possibility that the way in which our use of a term is causally regulated might develop as we get to know more about G and as our epistemic situation improves in other respects.

The requirement that an account of what a moral term refers to should be able to explain the reference-fixing characteristics has implications for the kind of non-moral property the term refers to. As a consequence, this requirement is also relevant to which type of moral theory specifies this non-moral property. As we saw above, on synthetic reductionism there should be explanations of reference-fixing characteristics in terms of the non-moral property that constitutes the reference of a moral term. For example, according to the view that 'right' refers to G, there should be explanations of reference-fixing characteristics such as 'People approve of right actions being performed' in terms of G. This means that G should consist of *that which* makes us—or at least some of us—have the beliefs and attitudes, and so forth, involved in these characteristics. This indicates that the theory which specifies what a moral term refers to is a normative moral theory rather than a meta-ethical theory. To see this, consider the following. It is often claimed that meta-ethical theories concern what moral terms mean. This indicates that such theories, among other things, are designed to systematise the features which function as conditions under which moral terms apply, where it is part of being linguistically competent to know that the terms apply to objects that have these features. As was suggested above, some reference-fixing characteristics can be understood to constitute such features, even if they are not necessarily had by what the

---

<sup>12</sup> For a more sophisticated account of causal regulation, see Boyd (1988), pp. 195, 209–212. See also Brink (2001), pp. 167–176; Copp (2000), pp. 113–120, and Sayre-McCord (1997a), pp. 269–271, 279–280.

terms applies to. However, meta-ethical theories are not designed to explain what it is about objects that make them have these features. Normative theories, on the other hand, can be understood to do that, among other things. If a normative theory is correct, it can be assumed to account for what it is about the objects to which a moral term applies that make us have the beliefs and attitudes, and so forth, involved in the reference-fixing characteristics. Accordingly, on synthetic reductionism, a moral term refers to a non-moral property of the sort involved in a normative theory.

There is a further reason to assume that what a moral term refers to on synthetic reductionism is specified by a normative theory. As mentioned above, on this view it is not part of being linguistically competent with respect to a moral term to know what the term refers to. It was also suggested that meta-ethical theories are, among other things, designed to systematise features which are such that it is part of being linguistically competent with respect to a moral term to know that it applies to objects that have these features. This suggests that, on synthetic reductionism, the reference of a moral term is not stated by a meta-ethical theory, but by a normative theory.

This view is also confirmed both by what synthetic reductionists say about their own position and the examples they give of moral theories that they think might determine the reference of moral terms.<sup>13</sup> Advocates of synthetic reductionism put forward this idea by saying that what a moral term refers to is determined by the best normative theory.

A third point concerns what sort of entities natural kind terms and moral terms, respectively, refer to. A natural kind term refers, unsurprisingly, to a natural kind. But it does not seem reasonable to assume that a moral term refers to a non-moral property that constitutes a natural kind—at least not if we understand such entities in the way Kripke and Putnam do. In particular, it does not seem plausible to assume that the non-moral property in question has the same fundamental causal role and explanatory function as a natural kind. Moreover, the various objects that have the non-moral property may fall into different natural kinds, with the result that the various instances of the non-moral property exemplify different natural kinds. But it can nevertheless be maintained that a moral

---

<sup>13</sup> See e.g. Brink (2001), p. 162; Copp (2000), p. 116; Merli (2002), p. 222; Railton (1989), pp. 157, 167–173, and Sayre-McCord (1997a), p. 286.

term refers to a non-moral property in a way that corresponds to how a natural kind term refers to a natural kind.<sup>14</sup>

In the last section, we saw that Kripke's and Putnam's account of the reference of natural kind terms can be examined by asking two questions that appeal to thought experiments. In the next section, I will examine synthetic reductionism by asking two analogous questions. But first I would like to comment on the relation between the open question argument and the kind of argument employed here. As mentioned in the introductory section, the procedure employed in the open question argument cannot be used to examine whether a certain term refers in the way outlined above. In inviting us to ask questions like 'Is whatever that is M really G?', the open question argument appeals to the knowledge we have *qua* linguistically competent persons of the meanings of certain terms or, more generally, of certain analytic truths. Appealing to this kind of knowledge, the argument thus employs our responses to appropriate questions to test whether certain terms have the same meaning. Since the view under discussion here is not based on meaning equivalence, it cannot be tested by appeal to this knowledge. However, the kind of argument with which we are concerned here resembles the open question argument in that both arguments rest on the idea that we, in virtue of being linguistically competent with respect to the terms at issue, have knowledge about the conditions under which the terms apply and hence respond to certain questions in a way reliably mirroring this knowledge.<sup>15</sup> As mentioned, the open question argument appeals to our knowledge of the meanings of terms or analytic truths. However, the present kind of argument appeals to our knowledge of something quite specific: whether a certain term refers to the stuff or property that causally regulates our use of the term and fulfils certain other requirements.<sup>16</sup>

---

<sup>14</sup> This point is stressed by Sayre-McCord (1997a), p. 271. See also Kim (1997), pp. 293–301, and Sayre-McCord (1997b), pp. 320–323.

<sup>15</sup> Cf. Horgan and Timmons (1992b), pp. 161–162.

<sup>16</sup> This difference between the open question argument and the kind of argument under consideration here means that the latter kind of argument is stronger than the open question argument in at least two respects.

First, the two kinds of arguments differ concerning the extent to which they have to be applied to various views about what a term means or refers to in order to yield significant results. As regards the procedure employed in the open question argument, we have to ask the appropriate questions about each single analysis and respond to the questions with doubt, to establish that there is no reductive analysis that states the meaning of the term at issue. (Unless, that is, we appeal to an inference to the

In the last chapter, I argued that doubt or lack thereof is the response that should be taken to be significant to questions put in the open question argument. Concerning the questions appealed to in the kind of argument with which we are concerned here, I have indicated that the appropriate responses are ‘yes’ or ‘no’. This difference is of no particular importance, especially as I use ‘lack of doubt’ and ‘doubt’ in such a way that they cover ‘yes’ and ‘no’. The reason why I have chosen to use different terms is merely an effect of the fact that in the kind of argument we are concerned with here, the responses seem to be made with less hesitancy, at least as regards natural kind terms. However, also as regards this kind of argument we may respond with weaker responses than an outright ‘yes’ or ‘no’. Concerning both kinds of arguments, strength in argument varies with strength in response.

#### 4. *The First Question*

In this section and the next, I will employ thought experiments in order to examine the plausibility of synthetic reductionism.<sup>17</sup> As in the thought

---

best explanation, as I did in the last chapter.) As regards the kind of argument we are concerned with here, we do not have to ask the two questions about each single view of what the term at issue refers to, to establish that it does not refer to a certain stuff or property in the way suggested. On the present view of reference, we only need to ask a question which indicates whether it is true that the term refers to the stuff or property that causally regulates our use of it and fulfils certain other appropriate requirements. The first question posed in this kind of argument is such a question, although there might be other considerations that yield the same result. If we respond to this question with ‘yes’ or lack of doubt, this indicates that the mentioned claim does not hold. This response thus suggests that the term does not refer to *any* stuff or property in the way proposed, whatever the stuff or property is.

Second, the open question argument and the kind of argument employed here differ as regards the significance of a response of lack of doubt or ‘yes’ to the questions put in them. As mentioned in the reply to objection (ii) in the last chapter, a response of lack of doubt to a question put in the open question argument cannot generally be taken to indicate that a certain truth is analytic. The reason is that there are presumably a number of truths that we do not doubt but that are not analytic. We run in other words the risk of taking a non-analytic truth for an analytic one. However, a response of lack of doubt or ‘yes’ to the second question—together with a negative answer to the first question—can be taken to indicate that a term refers to a stuff or property in the indicated way. One reason is that the latter argument does not appeal to knowledge of various analytic truths. Rather, it appeals to knowledge of something quite specific: whether a term refers to the stuff or property that causally regulates our use of it and fulfils certain other appropriate requirements. In the latter case, there seem to be no corresponding risk of confusion.

<sup>17</sup> Synthetic reductionism, or issues in relation to it, is critically discussed by e.g. Ball (1991), pp. 8–15; Barnett (2001), pp. 243–253; Blackburn (1993 (1990)), pp. 203–204; Blackburn (1998), pp. 119–121; Brandt (1996), chap. 6; Gampel (1996), pp. 191–209; Gampel (1997), pp. 147–163; Hare (1999 (1996)), pp. 81–86; Holland (2001), pp. 177–195; Horgan and Timmons (1992b), pp. 153–175, and other works by Horgan and

experiments in relation to ‘water’, we will imagine two kinds of possible worlds and in relation to these ask two questions. Analogous to our response to the first question as regards ‘water’, our response to the first question here indicates whether, for a non-moral property G that causally regulates our use of ‘M’, the following implication holds: It is synthetically necessary that, for any object x, if x is M, x is G. That is, the response to the first question indicates whether, in all possible worlds, G is a necessary condition for something to have M. However, we will also see that there are other considerations in connection with this question that might be used to determine the correctness of synthetic reductionism. As above, I will use ‘right’ as the example of a moral term.

In order to be able to ask the first question, we need to start by describing a scenario that complies with the outline of synthetic reductionism in the last section. Suppose that there are various characteristics that we associate with the term ‘right’ and that actions to which we apply the term typically have. Some of these characteristics are such that knowing that ‘right’ applies to actions that have them is considered to be part of what it means to be linguistically competent with respect to the term. However, these characteristics are not analytically necessary, something which is indicated by the fact that they do not pass the test provided by the open question argument. Rather, there is a non-moral property that causally regulates our use of the term by means of these characteristics. We may also assume that this property fulfils the other appropriate requirements hinted at in the last section. Now, suppose that according to the normative theory U this property is maximising happiness. On synthetic reductionism, ‘right’ refers to maximising happiness and rightness is identical with this property. According to this view, it is thus synthetically necessary that an action is right if and only if it maximises happiness.

Let us now imagine a possible world that resembles our world in many respects, but also differs from it in certain respects. In that world, a certain term ‘right’, phonetically and orthographically identical with our term, is used. The inhabitants of that possible world associate certain characteristics with this term and apply it to actions that have these

---

Timmons (see references below); Jackson (1998), pp. 144–150; Kim (1997), 293–230; Smith (1994), pp. 28–34; Sosa (1997), pp. 303–312; Tännsjö (1990), pp. 29–31, and Wong (1984), pp. 52–59.



characteristics. Some of these characteristics are such that it is considered to be part of being linguistically competent with respect to this term to know that it applies to actions that have them. These characteristics are generally the same as the characteristics that play the corresponding role in our world. Since these characteristics are generally the same, the inhabitants of the world under consideration have generally the same beliefs about what they apply their term 'right' to as we have about what we apply 'right' to. This is especially the case since both in our world and in their world these characteristics are 'common knowledge', something which is emphasised by the fact that some of these characteristics are connected in the mentioned way to linguistic competence. Consequently, they believe that people approve of the actions to which 'right' applies being performed, that people think it is of great importance that actions to which 'right' applies are performed, and so forth as regards the other characteristics. Moreover, since these characteristics are generally the same as those we employ, the inhabitants of that possible world to a great extent apply 'right' to the same kind of actions to which we apply 'right'. As in our world there is a non-moral property that causally regulates people's use of 'right' by means of the mentioned characteristics. However, in that world, the property that causally regulates people's use of 'right' is not the same as in our world. In the world under consideration, the property in question is specified by the normative theory DD. According to DD, the property consists in being desired to be desired. The explanation for why this property causally regulates their use of 'right' in virtually the same way as maximising happiness causally regulates our use of 'right', can be found in a certain morally neutral difference between our world and theirs. For example, we and they might differ slightly genetically, with the effect that we and they differ somewhat in temperament.<sup>18</sup>

We may now ask the first question: if an action in the possible world described above is such that it is desired to be desired, is it right? I think we are inclined to answer that it *might* be. At least, I do not think we are inclined to answer a straightforward 'no'.<sup>19</sup> This provides some reason to believe that the non-moral property that causally regulates our use of 'right' is not a necessary condition for an action to be right in that possible world.

---

<sup>18</sup> Cf. Horgan and Timmons (1992b), p. 165.

<sup>19</sup> Cf. Horgan and Timmons (1992b), p. 166.

This would in turn imply that the synthetically necessary biconditional mentioned above does not hold. Our response to the first question seems to differ when it is asked about ‘water’ and when it is asked about ‘right’. As far as ‘water’ is concerned, our response indicates that for something to be water in another possible world, it has to be the stuff that causally regulates our use of ‘water’, viz. H<sub>2</sub>O. However, our response to the first question as regards ‘right’ does not seem to provide support to the corresponding claim. Thus, whereas our response to the first question provides support to Kripke’s and Putnam’s view of the reference of ‘water’, it does not seem to support the view that ‘right’ refers in the way proposed by synthetic reductionism.

However, there are other, and perhaps more compelling, considerations in relation to the thought experiment put forward above that indicate that synthetic reductionism is mistaken. Consider the following. Suppose that we somehow were to communicate with the inhabitants of the possible world described above. Suppose further that we then come to know that they assert the following sentence: ‘It’s right to execute murderers’. Let us assume that we do not assert the corresponding sentence; that is, we do not believe that it is right to execute murderers. According to synthetic reductionism, we do not disagree with them but merely talk past each other. The reason is that since their use of ‘right’ is causally regulated by another non-moral property than our use of ‘right’, they do not refer to rightness. As Terry Horgan and Mark Timmons argue in their influential papers on ‘moral twin earths’, this description of the situation seems mistaken.<sup>20</sup> We have the intuition that we do disagree with the inhabitants of the possible world as to whether it is right to execute murderers. Hence, there is reason to believe that we and they refer to the same property: rightness.

---

<sup>20</sup> For this point see e.g. Horgan and Timmons (1992b), p. 165. Horgan and Timmons argues against synthetic reductionism and related views in the following works: Horgan and Timmons (1991), pp. 447–465; Horgan and Timmons (1992a), 221–260; Horgan and Timmons (1992b), pp. 152–175; Horgan and Timmons (1996), pp. 3–39; Timmons (1990), pp. 97–129, and Timmons (1999), chap. 2. Before Horgan and Timmons, Robert Merrihew Adams has discussed an argument based on a similar thought experiment: Adams (1981 (1979)), pp. 117–118. The works by Horgan and Timmons have provoked several critical responses; see e.g. Brink (2001), pp. 154–176; Copp (2000), pp. 112–137; Kraemer (1991), pp. 467–472; Merli (2002), pp. 208–240; Miller (2003), pp. 163–168; Sayre-McCord (1997a), pp. 267–292, and Zangwill (1997), pp. 509–518. Horgan and Timmons have replied to Copp in Horgan and Timmons (2000), pp. 142–149.

This contention is strengthened by a further reflection. Were we somehow to communicate with the inhabitants of mentioned possible world, it seems that we could have meaningful discussions with them about various issues pertaining to moral rightness. For example, it seems that we could debate with them in a rational way as to whether executing murderers is right or not. However, if we and they do not refer to the same property, it is difficult to see how this could be the case. As we and they would not disagree, but merely talk past each other, our debates with them would presumably be rather pointless.

Our considerations in relation to disagreement indicate that our intuitions as regards 'right' are not the same as regards 'water'. As we saw above, regarding the latter term we feel that the inhabitants of a possible world whose use of their term 'water' is causally regulated by something else than H<sub>2</sub>O refer to something else than we do with our term 'water'. However, as regards 'right' we are not of the corresponding view. Hence, neither in this respect does our intuition provide support to synthetic reductionism; rather, it indicates that this view is mistaken.

Moreover, it seems plausible to assume that we and the inhabitants of the possible described above world disagree about what the reasons are for actions being right. According to us, actions are right because they maximise happiness, but according to them, actions are right because they are desired to be desired. One way to express this is that we have another view of what *makes* actions right than they have. It is also reasonable to suppose that this difference explains why we disagree with them as to whether certain actions are right. (I will return to this issue in the next chapter.)

A similar reasoning holds concerning agreement about moral issues. Suppose that we come to know that the inhabitants of the mentioned possible world assert the following sentence 'It's right to help people in need'. Let us also assume that we assert the corresponding sentence; that is, we believe that it is right to help people in need. According to synthetic reductionism, in asserting this sentence, we do not really agree with them that it is right to help people in need. The reason is, again, that since their use of 'right' is causally regulated by another property than causally regulates our use of 'right', they do not refer to rightness. However, we seem to have the intuition that we and the inhabitants of that possible world do agree that

it is right to help people in need. Hence, there is reason to believe that they refer to the same property as we do. Moreover, while we agree with them in the indicated respect, it seems reasonable to assume that we and they have different views about what makes actions right.

These considerations suggest that a term—such as the term used by the inhabitants of the possible world described above—might refer to rightness even if its use is not causally regulated by the non-moral property that causally regulates our use of ‘right’. Hence, there is reason to believe that the terms refer to the same moral property. This indicates that ‘right’ does not refer to a non-moral property in the way synthetic reductionism maintains.

Let us return for a moment to the possible world described above. It might be argued that the normative theory embraced in that world is mistaken. As a result, it might be argued that we should not accept that, in that possible world, if an action is desired to be desired, it is right. The conception hinted at above is compatible with this view. According to this conception, both we and the inhabitants of the possible refer to rightness. However, it might be maintained that they are mistaken about what makes actions right and hence that the mentioned implication does not hold. On the other hand, it might be maintained that we are not thus mistaken. This conception seems to find some confirmation if we consider a person who embraces a particular normative conception. Consider for example a convinced utilitarian. It is reasonable to suppose that she would consider herself to be disagreeing with the inhabitants of the possible world as regards the normative conception they accept. One reason for this assumption is that she considers herself to be disagreeing with proponents of the corresponding normative conception in our world, just as she disagrees with proponents of virtue ethics and various deontological theories in our world. Hence, it is reasonable to assume that she would believe that the inhabitants of that possible world refer to rightness with their term ‘right’. However, she would most certainly maintain that they are mistaken about what makes actions right.

Although I have merely argued that synthetic reductionism is mistaken as regards one term, ‘right’, there seems to be no reason to believe that it would be more successful regarding other thin moral property terms. Thus, I think the arguments above suggest that synthetic reductionism is

mistaken also when it comes to these terms.

To argue against synthetic reductionism by means of the kind of thought experiment that is employed in the first question is quite expected and not very original. After all, Kripke and Putnam make use of similar thought experiments to establish their account of natural kind terms. Moreover, Horgan and Timmons appeal to one such thought experiment to refute synthetic reductionism. They argue, as I have done above, that this view is mistaken because people in a possible world whose use of a term is causally regulated by another property than our use of a corresponding term may disagree with us, not just talk past us, as synthetic reductionism implies. However, they never in effect ask the second question that I will consider in the next section. In the next chapter, I will argue that this is significant because if we take into account our considerations in relation to both questions, we will see that they provide support to realism. Horgan's and Timmons's argument against synthetic reductionism has evoked a number of interesting responses.<sup>21</sup> A thorough examination of synthetic reductionism would have to address the issues these responses raise. I will not do so here, however. As just indicated, my strategy will instead be to argue that these responses are better accounted for by realism than by reductionism.

### 5. *The Second Question*

Let us turn to the second question. Analogous to our response to the second question as regards 'water', our response to this question would indicate whether an implication of the following kind holds: It is synthetically necessary that, for any object *x*, if *x* is *G*, *x* is *M*, where *G* is a non-moral property that causally regulates our use of 'M'. That is, the response to the second question indicates whether, in all possible worlds, *G* is a sufficient condition for something to have *M*.

In order to be able to ask the second question, we should start with imagining a scenario that complies with the outline of synthetic reductionism in section 3. So assume that the scenario described in the last section holds. That is, suppose that our use of 'right' is causally regulated by the non-moral property specified by the normative theory *U*, i.e.

---

<sup>21</sup> See especially Brink (2001), pp. 154–176; Copp (2000), pp. 112–137, and Sayre-McCord (1997a), pp. 267–292.

maximising happiness.

Let us now imagine a possible world that resembles our world in certain respects, but also differs from it in certain other respects. Consider a world where there are actions that maximise happiness. However, in that world actions with this non-moral property typically do not have the characteristics we associate with 'right'. That is, in that world these actions do not generally evoke the beliefs and attitudes involved in these characteristics; they are not such that people approve of them being performed, etc. We may also assume that the inhabitants of that world have a term 'right' which is phonetically and orthographically identical with our term 'right'. They associate the mentioned characteristics with this term and apply it to actions that have them. But in that world, these actions typically are not such that they maximise happiness. Accordingly, their use of the term 'right' is not causally regulated by this property.

Let us now ask the second question: if an action maximises happiness in that world, is it right? I think we are inclined to answer that it *might* be. We would say that the action might be right in spite of the fact that these actions do not have the mentioned characteristics. Here our response resembles our response to the second question as regards 'water', although we in the present case presumably are less inclined to answer a straightforward 'yes'. Analogous to how we reasoned in relation to 'water', we may support our response by saying that an action in that world which maximises happiness is right, but that the inhabitants of that world do not recognise this. Hence, we have not found reason to believe that the non-moral property that causally regulates our use of 'right' is not sufficient for an action to be right in that world.

For there to be a possible world of the kind appealed to in the thought experiment above, it has to differ from our world, and this difference should explain why actions which have the non-moral property that causally regulates our use of 'right' do not have the mentioned characteristics in that possible world. That is, the difference should explain why, in that world, these actions are not such that people approve of them being performed, etc. Now, I think our response to the second question is reinforced in consideration of such differences. Return to the scenario according to which our use of 'right' is causally regulated by maximising happiness. Suppose, by way of illustration, that the difference which

provides the required explanation is that the inhabitants of the other possible world, due to their genetic traits, experience happiness very seldom. As a matter of fact, most of them have not experienced it at all, and if they have, they have done so only a few times. I do not think this explanation would have us to respond differently to the second question. Moreover, I do not think we would respond in any other way if we imagine possible worlds where the difference which provides the required explanation is any other than the one suggested. Imagine worlds that are like the world mentioned above, except that the relevant difference is any of these: due to a past event, the inhabitants feel guilt about happiness; they suffer from severe depression that makes them uninterested in happiness; they suffer from a certain disease that makes them experience happiness very seldom; that world is polluted in a certain way with the same effect. About each of these possible worlds, I think we are inclined to respond in the same way: we do not want to deny that an action that has the non-moral property which causally regulates our use of 'right' is right, in spite of the fact that it does not have the mentioned characteristics. To generalise, I think it is difficult to come to think of any difference that would make us deny this.

These considerations may also be seen in the light of what was said in the last section. When the first question was discussed there, it was suggested that the non-moral property which causally regulates our use of 'right' might not provide a necessary condition for an action to be right in another possible world. In particular, we saw that there are reasons to believe that a moral term does not refer to a non-moral property in the way maintained by synthetic reductionism. However, we do not seem to have corresponding results as regards the second question. It does not seem plausible to claim that an action which maximises happiness is not right in a possible world on the ground that happiness in that world, because of people's genetic traits, is experienced seldom or that happiness in that world, because of past events, is associated with guilt, etc. As was suggested above, it seems more plausible to say that these features may have the effect that the inhabitants of these possible worlds are prevented from *recognising* what is right. Thus, none of these differences provides reasons to assume that the non-moral property which causally regulates our use of 'right' does not provide a sufficient condition for an action to be right in these worlds. Moreover, I find it hard to come to think of any difference that

convincingly would show this.

What our response to the second question suggests is that given that a non-moral property causally regulates our use of ‘right’, an action in another possible world which has that property is right. This response indicates that as regards ‘right’, the implication mentioned above does hold. In line with what was said in the last section, it can be suggested that such a non-moral property *makes* actions right. (I will return to this proposal in the next chapter.) Although I have only argued for this conclusion concerning one moral term, ‘right’, I see no reason to believe that it does not hold also for other terms of the relevant kind.

## 6. *Concluding Remarks*

Advocates of synthetic reductionism maintain that a term ‘M’, applying to objects with a moral property M, and a term ‘G’, applying to objects with a non-moral property G, refer to the same property and that M and G consequently are identical. They are then committed to the truth of a synthetically necessary biconditional which provides a necessary condition for M and G being identical. In this chapter, I have argued that our response to the first question posed above gives some reason to believe that it is not synthetically necessary that if something is M, it is G. This would imply that the mentioned biconditional does not hold. But I have especially maintained that a moral term ‘M’ does not refer to a non-moral property G in the way proposed by synthetic reductionism. There are consequently reasons to believe that synthetic reductionism is mistaken. However, I have also suggested that our response to the second question indicates that it is synthetically necessary that if something is G, it is M. We have also seen some indications of the way in which non-moral properties can make objects have a moral property.

The arguments proposed in this chapter might be responded to in various ways.<sup>22</sup> For example, it might be argued that, on closer inspection, we do not have the intuitions appealed to in the arguments, or that these intuitions can be explained in a way that is consistent with synthetic reductionism. Advocates of this view might also try to characterise synthetic reductionism in such a way that it is consistent with our responses. As mentioned earlier, Horgan’s and Timmons’s argument against synthetic

---

<sup>22</sup> For various possible responses, see Horgan and Timmons (1992b), pp. 168–169.



reductionism has evoked a number of interesting defences of this view. Instead of considering these proposals, I will in the next chapter argue that realism is able to account for our considerations in relation to the two questions in a way superior to synthetic reductionism. Moreover, since the arguments put forward against synthetic reductionism in this chapter might be objected to, there will be reasons to return to this view in subsequent chapters.

## Chapter 4

### Realism and the Realist Formula

#### 1. Introduction

In the previous two chapters, I argued that there are reasons to believe that analytic reductionism and synthetic reductionism are mistaken. On the assumption that they comprise the two main forms of reductionism, these arguments give reasons to conclude that reductionism is mistaken. Some meta-ethicists would presumably take this result as an indication of the correctness either of non-cognitivism or error-theory. However, in chapter 1 we saw that there are *prima facie* reasons to reject both these views and instead adopt their respective contraries: cognitivism and success-theory. According to the classification of meta-ethical positions provided in that chapter, there are two kinds of meta-ethical positions which deny non-cognitivism and error-theory and affirm cognitivism and success-theory: reductionism and realism. Taken together, these points provide support to realism. That is, the arguments indicating that reductionism is mistaken, together with the arguments indicating the implausibility of non-cognitivism and error-theory, provide support to realism. However, we have also seen that the arguments against analytic and synthetic reductionism may be questioned. Moreover, although non-cognitivism and error-theory seem to have rather serious shortcomings, there are well-known arguments in support of these views. Thus, we have reasons to return to the two forms of reductionism and to consider the arguments put forward in favour of non-cognitivism and error-theory.

In this chapter, I will argue that our considerations in relation to the questions posed in the open question argument and the two questions posed in the last chapter provide further support to realism. More exactly, I will maintain that these considerations give reason to hypothesise that objects have moral properties in virtue of having non-moral properties in a manner I will refer to as 'the realist formula'. By suggesting the realist formula, these considerations confirm realism because this formula implies that moral properties are not reducible to non-moral properties.

In the next section, I argue that our reflections concerning the

mentioned questions suggest the realist formula, and hence provide support to realism, in the indicated way. In section 3, I argue that realists can make use of several of the ideas in synthetic reductionism.

## 2. *The Realist Formula*

As mentioned in the previous chapter, although there is no analytically necessary biconditional of the sort maintained by analytic reductionism, there are certain analytically necessary conditions that have to be fulfilled for moral terms to apply. I also mentioned the condition that seems most directly relevant to the relation between moral and non-moral properties: it is analytically necessary that an object has a moral property *M* *in virtue of* some of its non-moral properties. This can also be put by saying that it is analytically necessary that an object's moral property *M* *depends* on some of its non-moral properties or that some of its non-moral properties *make* it have *M*.<sup>1</sup> I will return to the importance of this relation between moral and non-moral properties in chapter 8. Now, the fact that this feature is an analytically necessary condition on the application of moral terms means that realists have to account for it. So even if realists are correct in claiming that moral properties are irreducible to non-moral properties, they have to account for the notion that objects have moral properties in virtue of having non-moral properties.

The analytically necessary condition on applications of moral terms, together with our considerations in relation to the questions posed in last two chapters, give reasons to hypothesise that the following formula accounts for the relation between moral properties and non-moral properties:

*The realist formula:* (i) It is analytically necessary that, for any object *x*, and for any moral property *M*, if *x* has *M*, then there is a set of non-moral properties *G* such that (A) *x* has *G*, and (B) it is synthetically necessary that, for any object *y*, if *y* has *G*, then *y* has *M*. (ii) *M* is not identical with any non-moral property.

The occurrence of analytic necessity at the opening of (i) is intended to capture the fact that it is analytically necessary that if an object has a moral

---

<sup>1</sup> Other ways to put this relation is to say that an object has *M* *because* it has certain non-moral properties or that certain non-moral properties *determine* that an object has *M*. One might distinguish between 'depend', 'determine', etc.; see Grimes (1991), p. 83, and Van Cleve (1990), p. 226. However, I will not do so here.

property M, it has it in virtue of some of its non-moral properties. The realist formula then says that if something has M, there is a set of non-moral properties G such that **(A)** the object has G. Thus understood, the realist formula says that if something has M, it has *some* set of non-moral properties, not that it has one particular set of non-moral properties which objects must have whenever they have M.<sup>2</sup> This claim provides some support from our responses to the questions posed in the open question argument and from our response to the first question posed in the last chapter: that there is no necessary implication from a moral property to a particular non-moral property. The realist formula says next, in **(B)**, that it is synthetically necessary that if an object has G, it has M. This claim reflects our response to the second question posed in the last chapter: that there is a synthetically necessary implication from a non-moral property to a moral property.<sup>3</sup>

What is said in **(i)** is compatible with a moral property M being identical with a non-moral property G, something that is irreconcilable with realism. However, the realist formula also includes **(ii)**: the claim that M is not identical with any non-moral property. The insertion of this claim into the realist formula is justified by our considerations in relation to the questions posed in the two preceding chapters. These considerations suggest that the meaning or reference of a moral term ‘M’ is not constituted by a particular non-moral property G in the way maintained by reductionism, i.e. a non-moral property with which M is identical. Accordingly, these considerations suggest that M is not identical with G.<sup>4</sup> Moreover, the

---

<sup>2</sup> In what follows, I will assume that the set of non-moral properties referred to in the realist formula is not empty. In so far as the set consists in more than one non-moral property, I will assume that it consists in a conjunction of non-moral properties. I will also assume that if a certain object has a moral property M, it has only one set of non-moral properties that makes it have M.

<sup>3</sup> As can easily be seen, the first part of the realist formula, **(i)**, is a variant of what is known as strong supervenience. In chapter 8, I will consider the relation between the realist formula and strong supervenience. For a version of strong supervenience which resembles the realist formula as to the way the two occurrences of ‘necessary’ is read, see Dreier (1992), p. 20. Cf. Danielsson (2001), p. 95, and McLaughlin (1995), p. 27. There are certain general problems that trouble any attempt to formulate a non-reductive dependence relation between properties; see e.g. Grimes (1988), pp. 152–160, and Kim (1993 (1990)), pp. 144–149. Unfortunately, I will have nothing to say about these difficulties here.

<sup>4</sup> According to one view of property identity, the truth of a necessary biconditional provides a necessary and sufficient condition for property identity. This means that if a property A and a property B are not identical, there is no true necessary biconditional involving A and B. On this view, **(ii)** in the realist formula—the claim that a moral property M is not identical with any non-moral property—implies that there is no true

insertion of **(ii)** can be motivated in terms of our responses to the questions posed in the open question argument and our considerations in relation to the first question posed in the last chapter, since these indicate that no necessary biconditional involving M and a non-moral property holds. If that is correct, In that case, a necessary condition for M being identical with a non-moral property is not fulfilled.<sup>5</sup>

As just indicated, on the realist formula there might be more than one set of non-moral properties in virtue of which objects have a moral property M. Put in a way inspired from the philosophy of mind: M is *multiply realisable*. In that case, more than one implication of the kind **(B)** is true.

In the light of what has been said above, the difference between reductionism and realism can be formulated in the following way. According to reductionism, there is a symmetric relation between a moral property M and a non-moral property G in that M is identical with G. According to realism, the relation between M and a set of non-moral properties G is asymmetric in that objects have M in virtue of having G, but M is not identical with G.

It might be argued that **(ii)** is insufficient to guarantee that there is an asymmetric relation of the mentioned kind. The reason is that according to some views on property identity, the fact that a property A is not identical with a property B is compatible with A and B being necessarily co-extensive, in which case a necessary biconditional involving A and B holds. However, in that case we would hesitate to claim that the relation between A and B is asymmetric. Consequently, on this view the realist formula is compatible with the truth of a necessary biconditional involving M and a non-moral property, and it might be argued that it therefore is insufficient to guarantee that there is an asymmetric relation of the kind mentioned above.<sup>6</sup> I have two comments on this argument. First, as we saw above there are some reasons to deny that there is any true necessary biconditional of this type; indeed, it is partly considerations to that effect that justify the insertion of **(ii)** into the realist formula. For that reason, the realist formula

---

necessary biconditional involving M and a non-moral property. However, this view on property identity is controversial. I will briefly return to these issues in chapter 8.

<sup>5</sup> For a version of strong supervenience that involves a condition similar to **(ii)**, see Depaul (1987), p. 433.

<sup>6</sup> Cf. Depaul (1987), pp. 433–434.

might be taken to suggest that no such necessary biconditional holds. Second, on the indicated view on property identity it might be the case that a moral property M and a non-moral property are not identical even if the mentioned type of necessary biconditional holds. So even if it would turn out that such a biconditional holds, the realist formula might be true.<sup>7</sup> However, in that case it might be incorrect to say that it states an asymmetric relation.

In subsequent chapters, I will say that, according to the realist formula a set of non-moral properties *makes* objects have a moral property. Thus, to formulate the realist formula in a simple way that I occasionally will make use of, it says that an object which has a moral property has a set of non-moral properties such that it makes objects have that moral property. (It should be admitted, however, that my usage of ‘make’ to a certain extent is a stipulation, since one might want to utilise this term even if one does not share this view of the relation between moral and non-moral properties.)

As we have seen, the realist formula incorporates our considerations in relation to the questions posed in the open question argument and the argument in the last chapter. It is therefore reasonable to hypothesise that the realist formula explains these considerations. It can in other words be assumed that we respond as we do to these questions because the relation between moral properties and non-moral properties is the one described in the realist formula.

One aspect of the explanatory capacity of the realist formula is particularly worth observing. In discussing the first question posed in the last chapter, we imagined a possible world where the inhabitants have their use of a term ‘right’ causally regulated by another non-moral property than causally regulates our use of ‘right’. In spite of this, we believe that we may disagree with these people as to whether actions are right. Hence, it seems

---

<sup>7</sup> Suppose, in accordance with the objection above, that there is a true necessary biconditional involving a moral property M and a non-moral property G, but that M and G are not identical. It might then be argued that the realist formula should be reformulated in such a way that it captures the fact that such a necessary biconditional holds. However, I do not think such a reformulation would be desirable. Even if such a necessary biconditional holds, this is not something which we want to commit ourselves when we assert that an object has a moral property M. In particular, we do not want to commit ourselves to the view that it is necessary that whatever object has M has one particular non-moral property which objects must have whenever they have M. The reason is that we want to leave open the possibility that there is more than one set of non-moral properties in virtue of which objects have M.

that we and they refer to the same property: rightness. We also saw that it is reasonable to say that our view of what makes actions right differs from theirs. The realist formula can explain this. On this conception, both we and the inhabitants of that possible world refer to rightness because both we and they refer to the property as it is characterised in the realist formula. Roughly put, in asserting that an action is right, both we and they claim that it has a set of non-moral properties such that it is synthetically necessary that whatever action has that set of non-moral properties is right. However, we and they have different views about what makes actions right because we and they have different views about what such a set of non-moral properties consists in. More precisely, we and they differ as to whether it holds about a certain set of non-moral properties that it is synthetically necessary that whatever action has that set of non-moral properties is right. As we saw above, we might believe that the inhabitants of the other possible worlds are mistaken about what makes actions right. Our differing views in this regard might explain why we disagree with them as to whether certain actions are right. (I will return to similar considerations in the next chapter.)

Moreover, the realist formula furnishes realists with an account of the way in which non-moral properties make objects have moral properties which implies that moral properties are not reducible to non-moral properties. Thus, our considerations in relation to the various questions posed earlier provide support to realism because they suggest a formula which states a relation between moral and non-moral properties which implies that the former are irreducible to the latter.

### *3. Reusing Synthetic Reductionism*

In the previous chapter, I argued that there are reasons to assume that synthetic reductionism is mistaken. This conclusion notwithstanding, I think a number of ideas in synthetic reductionism can be incorporated in a realist account of moral properties. In this section, I provide a bare sketch of such an account. Since its essential parts have been mentioned in the preceding chapter, I will make it brief. The idea, roughly put, is that what according to synthetic reduction holds for the non-moral property that constitutes the reference of a moral term holds according to realism for a set of non-moral properties that makes objects have a moral property. Once more, I will use 'right' by way of illustration, but I believe that this account is valid for other

thin moral properties as well.

In the previous chapter, I suggested that there are some analytically necessary conditions that have to be fulfilled for 'right' to apply. (In the last section, we saw the relevance of one of them.) However, the open question argument indicates that these conditions are not sufficient for us to be able to apply the term correctly. To explain how this is possible, it may be hypothesised that there are certain characteristics that work analogously to how reference-fixing characteristics work concerning natural kind terms. We associate such characteristics with 'right' and they are typically had by the actions to which we apply the term. By means of these characteristics, we are able to apply 'right' correctly even if we are not aware of the correct account of the non-moral properties that make actions right. Some of these characteristics may be such that knowing that 'right' applies to actions that have them is part of what it means to be linguistically competent in respect of the term. However, unlike analytically necessary conditions, they are not necessarily had by what the term applies to.<sup>8</sup> Rather, these characteristics can fulfil this function because actions that are right have certain non-moral properties which are causally responsible for these characteristics and so causally regulate our use of the term via these characteristics. However, there might be other factors than the reference-fixing characteristics that help us to apply 'right' correctly; for example, it might be the case that we are aware of some aspects of what makes actions have that moral property.

Suppose that some actions have a set of non-moral properties *G* that causally regulates our use of 'right' and that fulfils the other appropriate requirements.<sup>9</sup> To see the implication of this, we should return to our response to the second question posed in the last chapter. This response suggests that it is synthetically necessary that if an action has *G*, it is right. In terms of the usage adopted above, *G* is a set of non-moral properties that makes actions right. This set of non-moral properties will consequently figure in the kind of implication (**B**).

In the preceding chapter, it was suggested that in order for a view to the effect that 'right' refers to a certain non-moral property to be justified, it

---

<sup>8</sup> In connection with realism, it might be misleading to call these characteristics '*reference-fixing characteristics*'. Perhaps they should rather be called '*dependence-fixing characteristics*' or something of the kind. I will however continue to use the first phrase so as to keep the connection to synthetic reductionism.

<sup>9</sup> If there are number of sets of non-moral properties of this kind, it might be the case that one of them only causally regulates a particular aspect of our use of the moral term.



has to be able to explain the reference-fixing characteristics and so how the property can causally regulate our use of the term. Similarly, in order for a view to the effect that a certain set of non-moral properties makes actions right to be justified, it has to be able to explain these characteristics. Synthetic reductionism and realism differ concerning what provides the required explanation. According to synthetic reductionism, it is the non-moral property to which 'right' refers that is causally responsible for the reference-fixing characteristics. According to realism, it is what makes actions right that is causally responsible for these characteristics. Thus, if G is a set of non-moral properties that makes actions right, G should help explaining the mentioned characteristics.

The demand that a view of what 'right' refers to should be able to explain the reference-fixing characteristics suggests, for reasons supplied in the last chapter, that the kind of theory that specifies the reference of the term is a normative theory. Similarly, the demand that a view of what makes actions right should provide such an explanation might be taken to indicate that it is a normative theory that specifies the right-making properties. However, as far as the realist formula is concerned, there might be a number of sets of non-moral properties that make actions right and these sets might be heterogeneous. Moreover, a set of non-moral properties that makes actions right might consist of a conjunction of a number of non-moral properties and these properties may be quite heterogeneous. Furthermore, these properties might be related to each other in various ways. It seems doubtful whether normative theories—understood as proper theoretical accounts that are intended to be systematic and coherent—are designed to specify such sets of non-moral properties. I will therefore put the present point in more general terms and say that what specifies such a set is a 'normative conception'.

What was said about epistemic method in relation to synthetic reductionism also applies to realism. It might be the case that a scientific or empirical method is unavailable, or plays a modest role, in the process of finding out what makes actions right. However, other methods which are not necessarily scientific or empirical, although some of them can be, might be pertinent. Some version of coherentism appears to be a promising candidate. Moreover, analogous to what was argued in the previous chapter, a set of non-moral properties that makes actions right does not constitute a

natural kind.

This sketch comprises various parameters that realists can understand in various ways. Accordingly, different versions of realism emerge depending on, among other things, how the following questions are answered: **(i)** What are the analytically necessary conditions that have to be fulfilled for ‘right’ to apply? **(ii)** What are the reference-fixing characteristics in respect of ‘right’? **(iii)** What makes actions right? **(iv)** What is the correct epistemic method to determine this? Corresponding questions can of course be asked about other moral terms and properties. With the exception of some remarks on **(i)** and **(iv)**, I will not take up these questions. This limitation is important, since the plausibility of realism might depend on how they are answered.

#### *4. Concluding Remarks*

In this chapter, I have argued that our considerations in relation to the questions posed in chapters 2 and 3, together with some arguments against non-cognitivism and error-theory, provide support to realism. Moreover, I have argued that these considerations give reasons to adopt the realist formula, which describes the way in which non-moral properties make objects have moral properties. In doing so, the mentioned considerations provide further support to realism, since the realist formula implies that moral properties are not reducible to non-moral properties. We have also seen that realism can make use of many of the insights of synthetic reductionism. However, as we have seen in the two previous chapters, the arguments put forward against reductionism can be questioned. Moreover, there are influential arguments in favour of non-cognitivism and error-theory. It should therefore be stressed that the case put forward in support of realism above is by no means conclusive. Rather, I regard realism as a promising hypothesis that is in need of further evidence and assessment. In the following three chapters, I will argue that realism, much owing to the realist formula, is able to account for certain meta-ethical issues and that realism therefore is preferable to reductionism, non-cognitivism and error-theory. It might however be argued that the realist formula is vulnerable to several objections. I postpone a discussion of these objections to chapter 8.

## Chapter 5

### Explaining Moral Disagreement

#### 1. *Introduction*

One of the features most commonly associated with morality by philosophical laymen is presumably disagreement. On a prevalent view, disagreement permeates morality and justifies far-reaching sceptical conclusions in regard to it. Among meta-ethicists there is controversy as to the philosophical significance of moral disagreement, but an influential line of thought follows the popular view. Proponents of non-cognitivism and error-theory have accordingly argued that reflections on the nature of moral disagreement should make us conclude that there are no moral properties. In this chapter, I will maintain that realism, in contrast to reductionism, is able to defend the existence of moral properties against such arguments. However, the occurrence of moral disagreements has given rise to a number of arguments, and I will comment only on those I find most troubling for the view that there are moral properties.<sup>1</sup>

The arguments I have in mind are generally based on the assumption that moral disagreements have some distinctive feature that gives reason to believe that there are no moral properties. In the next section, I give an outline of such an argument which I will refer to as ‘the argument from moral disagreement’. As I construe this argument, it contains a placeholder where different features of the mentioned kind can be inserted, thereby generating different versions of it. Two features of moral disagreements are typically appealed to: their being widespread or persistent. After noticing in section 3 that there are various types of moral disagreements, I discuss in the rest of the chapter the premises of the argument in turn, with some exception. In section 4, I consider the relation between moral disagreement and error. In section 5, I argue that the version of the argument from moral disagreement according to which the distinctive feature of moral disagreements consists in their being widespread is unconvincing. In section 6, I maintain that a more cogent version of the argument is obtained if this

---

<sup>1</sup> For a systematic discussion of the meta-ethical significance of moral disagreement, see Tersman (2002), chap. 2–5.

feature is assumed to consist in moral disagreements being persistent. The basic reason is that the persistence of moral disagreements is closely connected to a phenomenon I will refer to as ‘normative divergence’, by which I mean, briefly put, that people may disagree even if they have different normative conceptions and therefore take the moral term in question to have different conditions for application. In sections 7 and 8, I argue that reductionism fails to explain persistent moral disagreements, mainly because it fails to account for normative divergence. In section 9, I argue that relativism does not provide any help to reductionism. In sections 10, 11 and 12, I argue that realism, by employing the realist formula, succeeds in accounting for normative divergence and is able to explain persistent moral disagreements. I conclude in the final section that this means that realism is able to reject the argument from moral disagreement.

## 2. *The Argument from Moral Disagreement*

J. L. Mackie is presumably the philosopher most closely associated with the view that the phenomenon of moral disagreement provides an argument against the existence of moral properties. In a much-quoted passage he writes:

The argument from relativity has as its premiss the well-known variation in moral codes from one society to another and from one period to another, and also the differences in moral beliefs between different groups and classes within a complex community. [- - -] [R]adical differences between first order moral judgements make it difficult to treat those judgements as apprehensions of objective truths. [- - -] [T]he argument from relativity has some force simply because the actual variations in the moral codes are more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values.<sup>2</sup>

Mackie does not state his argument very clearly, and my purpose is not to offer an interpretation of his contention. However, Mackie’s reasoning can be understood to exemplify the argument from moral disagreement. This argument can be represented as follows:<sup>3</sup>

---

<sup>2</sup> Mackie (1977), pp. 36–37.

<sup>3</sup> Similar ways of reasoning are found in e.g. Loeb (1998), pp. 282–283; Tersman (2002), esp. chap. 2; Wright (1992), pp. 150–157, and Wright (1996), pp. 13–14. That the argument should be understood in roughly this way is agreed to by those who oppose it; see e.g. Brink (1989), pp. 197–198; Shafer-Landau (2003), pp. 215–216, and Wreen (1985), pp. 149–150. Sometimes arguments similar to the argument from moral disagreement are assumed to support relativism rather than the view that there are no moral properties; see e.g. Harman in Harman and Thomson (1996), pp. 8–14, and Wong

- (1) If there are any moral properties, the best explanations of moral disagreements imply that there is an error in relation to a moral property.
- (2) Moral disagreements have a distinctive feature F.
- (3) Due to feature F, the best explanations of such disagreements do not imply that there is an error in relation to a moral property.
- (4) Therefore, there are no moral properties.

Different versions of the argument are obtained by letting ‘F’ in premise (2) represent different features of moral disagreements.<sup>4</sup> Proponents of the argument from moral disagreement often take this feature to consist in such disagreements being widespread. According to this version of the argument, the fact that moral disagreements are widespread makes it unlikely that they are best explained by people having, as Mackie formulates it, ‘seriously inadequate and badly distorted’ ‘perceptions’ of a moral property. It is simply regarded as implausible that people are in error to such a great extent such an explanation would require. Sometimes the mentioned feature of moral disagreement is instead supposed to consist in their being persistent. This version of the argument gains support from the idea that since these disagreements concern issues that have been extensively debated without any solution having emerged, they are not best explained in terms of an error in relation to a moral property. It is simply thought that if they were best explained in this way, this is something that would have been found out, considering that these issues have been persistently debated.

There are other features of moral disagreements than their being widespread or persistent that can be appealed to in the argument from moral disagreement, especially what I will call ‘normative divergence’. However, since these two features are most commonly thought to be operative in the argument, it seems reasonable to focus on versions of it that take their point of departure in them. Proceeding in this way will also highlight significant interconnections between various aspects of moral disagreements. This will

---

(1984), esp. chap. 8. Cf. Sayre-McCord (1991), pp. 163–172. I will comment briefly on relativism in section 9.

<sup>4</sup> Another argument is that there are no moral properties because if there were, people would agree about moral issues to a much larger extent than they do. (For a discussion of an argument on that line, see Wiggins (1987 (1983)), pp. 162–171.) One difficulty with this argument is that moral agreements might have various explanations that do not involve the assumption that there are moral properties. Cf. Williams (1985), p. 136.

be especially clear in section 6 where the close connection between the persistence of such disagreements and normative divergence will become evident.

The conclusion of the argument from moral disagreement entails that reductionism and realism is false. To avoid it, advocates of these views have to deny some of its premises. The conclusion can be taken to support either error-theory or non-cognitivism, depending on whether moral sentences have truth-value or not.

As the argument of disagreement has been construed here, it focuses on the conclusion that there are no moral properties. However, the argument also implies that the best explanations of moral disagreements—if there are any explanations of that kind—are provided in terms of something other than an error in relation to a moral property. Some meta-ethicists, of which Mackie is an example, focus on this aspect of the argument rather than the one accentuated here. Accordingly, advocates of the argument propose explanations of moral disagreements which they believe are superior to those implying the existence of a moral property. As can be seen in the quotation above, Mackie claims to do so by combining his error-theory with the idea that moral disagreements reflect different ‘ways of life’. Non-cognitivists typically argue that moral disagreements are best explainable in terms of people having conflicting non-cognitive states.<sup>5</sup>

### 3. *Types of Moral Disagreement*

In the meta-ethical literature moral disagreement is usually considered as a homogeneous phenomenon. I think it is worthwhile, however, to distinguish between different types of such disagreements. One reason is that the kind of argument outlined in the last section varies in strength depending on which type that is under discussion. Another is that the plausibility of certain responses to it varies accordingly. These reasons will become clearer as we proceed.

The field of moral disagreement can of course be divided in a number of different ways. For my purposes it is expedient to make a tripartite division based on *about what* people disagree. The resulting distinctions can be combined with the distinction between widespread and persistent

---

<sup>5</sup> For non-cognitivist accounts of moral disagreement, see e.g. Nowell-Smith (1954), pp. 193–197, and Stevenson (1963), essay I, II and IV. See also references to Hare below.

disagreements.

(i) The most basic type of moral disagreement occurs when people disagree as to whether a *particular object* has a certain moral property M or not.

(ii) Another type of moral disagreement occurs when people disagree as to whether *objects belonging to a certain kind*—e.g. abortion, bigamy, death penalty, eating meat, euthanasia and raising animals for food—have a certain moral property M or not. Here we find the cases most often appealed to in the argument from moral disagreement.

(iii) A third type of disagreements occurs when people disagree as to whether a certain *normative conception* as regards a moral property M holds or not. In this kind of disagreement, people have different views about which normatively relevant non-moral properties objects should have in order to have M.<sup>6</sup> Sometimes this is put by saying that they have different views about which the good-making or right-making properties are. Formulated in another way, they have different views about which, in terms of normatively relevant non-moral properties, the conditions are for ‘M’ to apply. Examples of disputes about normative conceptions are debates about normative theories, such as those between utilitarians and deontologists. However, all disputes about normative conceptions do not concern views that are intended to be so systematic and coherent that they properly can be called normative theories; hence, the term ‘normative conception’ seems more to the point.

#### 4. *Moral Disagreement and Error*

In the remaining sections of this chapter, I will consider the premises of the argument from moral disagreement in turn, with some exception. Thus, we start with premise (1). One way to respond to the argument is to refute this premise. However, I do not think this response is plausible.<sup>7</sup>

Let us start by observing that on the assumption that there are moral properties, a moral disagreement involves an error in relation to a moral property. So assume that there are moral properties. It is then reasonable to suppose that declarative sentences pertaining to these properties are either true or false. A disagreement occurs in case someone asserts a sentence and

---

<sup>6</sup> Unfortunately, this is rather vague. By saying that the non-moral properties at issue are ‘normatively relevant’, I want to indicate that the views in question are not meta-ethical.

<sup>7</sup> In section 9, I will consider relativism, which can be understood to deny premise (1).

someone else denies it. It follows that if two persons disagree in any of the ways mentioned in the previous section, some of them claims something that is false. On the assumption that there are moral properties, a person is then in some sort of error in relation to a moral property that explains this and so explains the disagreement.<sup>8</sup> (What an ‘error in relation to a moral property’ amounts to will become clearer as we proceed.) This being the case, it seems plausible to claim that if there are moral properties, any reasonable explanation of a moral disagreement involves the assumption that there is an error of the mentioned kind.<sup>9</sup> Admittedly, this might be an error which is very difficult, perhaps even impossible, to avoid. Even so, the person in question is in error in the sense that there is some consideration in relation to a moral property such that *if* she somehow had differed with respect to that consideration, she would not have claimed something false. In the subsequent sections, I will consider various proposals of such considerations.<sup>10</sup>

Explanations of moral disagreements may involve further aspects in addition to the assumption that there is an error in relation to a moral property. Thus, on the view that there are moral properties, it seems more correct to say that reasonable explanations of such disagreements *imply* that there is an error of the mentioned sort, rather than that they outright *consist* in the stating of such an error. These additional aspects of explanations might be felt to be more important than the error they make reference to,

---

<sup>8</sup> This should not be taken to imply that the person who asserts a true sentence is not subject to any kind of error. Indeed, she might have reached her opinion in a defective way. For this reason it might perhaps be more accurate to say that the proper explanation of a moral disagreement implies that *at least* one of the persons involved in it is mistaken. Correspondingly, it might be more accurate to say that there is at least one error. However, in what follows I will ignore this complication.

<sup>9</sup> This view is in accordance with Crispin Wright’s principle of ‘cognitive command’. Wright takes this principle to imply that ‘one obligation of the moral realist will be to hold, and therefore to justify holding, that moral disagreements [. . .] have to involve defects of process or materials: at least one of the protagonists has to be guilty of a deficiency in the way he arrives at his view, or to be somehow constitutionally unfit’ (Wright (1996), p. 14). See also Wright (1992), pp. 92–93, 143–157. Wright may also be understood to endorse premise (1).

<sup>10</sup> It should be noted that even if error-theory accepts that people might be in error because the moral sentences they assert are false, it would be misleading on this view to say that they are in error *in relation to* a moral property, as there is no moral property to be in error in relation to. Rather, on this view people are in error in so far as they assume that there are any moral properties. Thus, error-theorists do not provide the kind of explanation proposed by advocates of the existence of moral properties. In particular, they do not consent to the view that *the best explanations* of moral disagreements imply that there is an error in relation to a moral property. This is indicated by Mackie’s appeal to explanations in terms of different ‘ways of life’.



and different explanations might vary in value because of them.<sup>11</sup> Accordingly, the explanations that are best may be so because they contain other aspects than the stating of the mentioned kind of error. However, from the claim that any reasonable explanation of a moral disagreement implies that there is an error in relation to a moral property, it follows that the *best* explanations of such disagreements imply that there is an error of that sort, even if they contain other aspects as well.

It might be objected that the above reasoning is defective because it ignores the possibility of moral terms being vague. On this objection, if a moral term 'M' is vague, there are cases where it is neither true nor false that an object has the moral property M.<sup>12</sup> This means that if someone claims that something has M and someone else denies this, it does not follow that any of them claims something that is false. Then there is no disagreement that is to be explained in terms of an error. I have three comments on this objection. First, according to one view of vagueness, when 'M' is vague, it is either true or false that an object has M; it is just that we cannot *know* whether this is the case or not. Second, even if the account of vagueness appealed to in the objection is correct, it is hardly the case that vagueness accounts for all, or even particularly many, cases of assumed moral disagreements. Hence, in so far as one believes there are moral properties, one has to explain such disagreements in terms of an error in relation to a moral property in some, and probably most, cases. Third, we can accommodate the mentioned view by reformulating (1) to say 'If there are any moral properties, the best explanations of moral disagreements imply that there is an error in relation to a moral property, *at least unless the pertinent terms are vague*'.<sup>13</sup> However, in view of the two first remarks, I will ignore this complication in what follows.

There is a further point that is relevant in the present context. In the introductory chapter, we saw that both non-cognitivism and error-theory have difficulties accounting for moral disagreement. According to non-

---

<sup>11</sup> For example, suppose it is proposed that people disagree morally because they embrace different religious convictions. This information might be thought to be more valuable than the information that there is some sort of error involved in the disagreement. However, on the present conception, the proposed explanation suggests that the reason why the people in question disagree is that their different religious convictions make some of them subject to an error of the indicated kind.

<sup>12</sup> See e.g. Shafer-Landau (1995), pp. 83–96.

<sup>13</sup> Cf. Wright (1992), pp. 144–145.

cognitivism, moral sentences do not have truth-value. This means that people who debate moral issues do not disagree in the sense that they assert contradictory moral sentences. This opposes our notion of moral disagreement. According to error-theory, moral sentences have truth-value. Consequently, people who are involved in moral disagreements do assert contradictory moral sentences. However, the participant in a disagreement who asserts a moral sentence in which a moral property is ascribed to an object is always wrong, and the explanation is simply that there are no moral properties at all, not that she is wrong about a particular moral issue. Also this opposes our notion of moral disagreement.<sup>14</sup> Advocates of reductionism and realism do not have these difficulties. Contrary to non-cognitivists, they claim that moral sentences have truth-value so that people who debate moral issues can genuinely disagree by asserting contradictory moral sentences. Contrary to error-theorists, they claim that there are moral properties so that people who assert moral sentences that ascribe moral properties to objects are not always necessarily wrong and wrong for an awkward reason. This indicates that there are *prima facie* reasons to maintain that the best explanations of moral disagreements are made on the assumption that there are moral properties. Consequently, *prima facie* it seems that the best explanations of moral disagreements imply that there is an error in relation to a moral property. I will return to this consideration in the last section.

##### 5. *Widespread Moral Disagreement*

We may now continue with considering premise (2) in the argument from moral disagreement. As mentioned above, feature F in that premise is regularly taken to consist in moral disagreements being widespread. It is often supposed that disagreement belonging to type (ii) have this feature. Moreover, it is typically assumed that there are widespread disagreements of this type between people belonging to different cultures.

However, I do not find this version of the argument from moral disagreement especially convincing.

One reason for this contention is that it is questionable whether moral disagreements are widespread in the relevant way. Let us first notice that the

---

<sup>14</sup> In chapter 1, I formulated this difficulty in terms of the person who is right, but it can be formulated just as well in terms of the person who is wrong.

claim that such disagreements are widespread involves a covert comparison with disagreements within other areas than morality. The claim that moral disagreements are widespread will only have force if this kind of disagreement is *more* widespread than disagreements within other areas. This is so since we presumably are not prepared to accept that the argument from moral disagreement applies generally to other areas. That is, we are not generally prepared to conclude that a certain kind of properties does not exist because there are disagreements in relation to it, even if these disagreements are widespread. If the argument ‘spread’ in this way to other areas, this would probably have us question the argument itself, not the existence of these properties.<sup>15</sup> Next it should be observed that the question of whether moral disagreements are more widespread is an empirical one that seems very hard to settle. Suppose the focus is on moral disagreements between persons belonging to different cultures. To determine whether such disagreements are more widespread would not only require thorough empirical investigations of the morality in different cultures and careful interpretations of the resulting data; it would also require comparisons with similar investigations in other areas to settle the relative amount of disagreement within the moral area. Until such an investigation has been carried out, something that is unlikely to happen, we are not in the position to tell whether the prevalence of moral disagreements can be legitimately appealed to in the argument from moral disagreement.

It might also be denied that moral disagreements are more widespread than disagreements in other areas. The claim that moral disagreements are widespread suggests a picture of morality where people have conflicting views about very many moral issues. However, this is hardly an impression the meta-ethical literature warrants.<sup>16</sup> Often authors just take for granted that moral disagreements are widespread. But in so far as they present any support of this view, they come up with quite short lists of types of moral issues about which there is supposed to be disagreement. Moreover, these lists tend to contain the same items.<sup>17</sup> Thus, rather than demonstrating that moral disagreements are widespread, the literature

---

<sup>15</sup> Cf. Shafer-Landau (2003), p. 220.

<sup>16</sup> Cf. Wreen (1985), pp. 149–150.

<sup>17</sup> For such a list, see e.g. Harman in Harman and Thomson (1996), pp. 8–11. Some authors argue that people within a certain culture—but also people belonging to different cultures—tend to reach moral agreement under favourable circumstances; see e.g. Brink (1989), pp. 208–209.

indicates that they are isolated phenomena that concern rather few but recurrent issues.

Furthermore, it seems possible to explain why moral disagreements might *appear* more widespread than disagreements in other areas, although they are not. Moral disagreements might receive more attention than other kinds of disagreement because they are more easily discernible, have practical implications and therefore are significant for our lives.<sup>18</sup> Disagreements in other areas may be as widespread but not as easily discernible or especially consequential. There is also another way in which moral disagreements may just appear more widespread. Suppose that a certain kind of actions, bigamy for example, is judged to be wrong in one culture but not in another. Then there appears to be a disagreement between these people as to the wrongness of these actions. However, this view can be questioned on the ground that people in these cultures may apply the same normative conception to different circumstances.<sup>19</sup> For example, in both cultures utilitarianism might be embraced. The circumstances in one of the cultures may be such that bigamy makes people happy, whereas the circumstances in the other culture may be such that it makes people unhappy. People in the first culture may then be correct in so far as bigamy in *their* culture, with *its* circumstances, is not wrong, whereas people in the second culture may be correct in so far as bigamy in *their* culture, with *its* circumstances, is wrong.

However, suppose that moral disagreements are more widespread than disagreements in other areas. In that case, a defender of the existence of moral properties can respond by denying premise (3). On realism, she may do so in basically two ways. First, she may argue that widespread disagreements can be explained in terms of an error as regards non-moral facts. Second, she may argue that they can be explained in the same way as persistent disagreements. I will discuss these types of explanations in sections 8 and 11, respectively.<sup>20</sup>

---

<sup>18</sup> Cf. Wellman (1975), p. 212.

<sup>19</sup> Cf. Brink (1989), pp. 200–202; Railton (1993), pp. 282–283; Ross (1939), pp. 18–19, and Wreen (1985), pp. 150–152. This kind of explanation is criticised by e.g. Mackie (1977), pp. 37–38.

<sup>20</sup> Above I mentioned that the version of the argument from moral disagreement under considerations typically appeal to the view that disagreements of type (ii) are widespread. However, it might be objected that disagreements of type (iii)—disagreements concerning normative conceptions—also may be widespread. Mackie can be taken to suggest this when he refers to ‘variations in moral codes’ in the quotation above. But if it

## 6. *Persistent Moral Disagreement and the Phenomenon of Normative Divergence*

In the last section, I argued that the version of the argument from moral disagreements that takes feature F in premise (2) to consist in moral disagreements being widespread is not particularly persuasive. We may continue with enquiring if a more powerful version of the argument is obtained if feature F is assumed to consist in moral disagreements being persistent.

The most obvious examples of persistent moral disagreements are disagreements of type (iii), i.e. disagreements about normative conceptions. In moral philosophy there are such disagreements as regards rightness between proponents of various normative theories such as, to name but a few, utilitarians and deontologists, various kinds of utilitarians, various kinds of deontologists, deontologists and virtue ethicists. However, not all disagreements about normative conceptions are disagreements about normative theories.

Persistent disagreements are not only to be found in disagreements of type (iii), but also in type (i) and (ii). However, it is reasonable to assume that persistent disagreements of type (i) and (ii) typically can be accounted for in terms of disagreements of type (iii). It seems in other words plausible to suppose that if people have debated some moral issue belonging to type (i) and (ii) extensively without having reached any solution, this is because they embrace conflicting moral conceptions. This underlying disagreement may thus account for why they have not resolved their disagreements at the other levels. Consider disagreements of type (i). It is not too uncommon that people disagree for example about the rightness of a particular action and after a lengthy discussion come to suspect that there is a more fundamental disagreement that explains the original one. Persistent disagreements of type (ii) are even more readily explainable in the indicated way. Consider for example the persistent debates about abortion, death

---

is difficult to show that disagreements of type (ii) are widespread, it seems at least as difficult to show that disagreements of type (iii) are widespread. The reason is that the only way to show that disagreements of the latter type are widespread seems to be to show that disagreements of the former type have this feature. That is to say, the prevalence of disagreements belonging to type (iii) will make it known through disagreements belonging to type (ii). However, in case disagreements of type (iii) are widespread, it might be suggested that they can be explained in the same way as persistent disagreements.

penalty, euthanasia, eating meat or raising animals for food. These issues have been debated extensively without agreement having been reached. It therefore seems plausible to assume that the underlying reasons for such disagreements regularly are to be found in conflicting normative conceptions. Indeed, quite often these issues are explicitly debated in such terms.<sup>21</sup>

Now, as R. M. Hare in particular has brought attention to, disagreements about normative conceptions have a remarkable feature.<sup>22</sup> This feature is relevant to the prospect of explaining persistent moral disagreements in terms of an error in relation to a moral property. In disagreements about normative conceptions, the antagonists have different views about the conditions under which a moral term applies. For example, one of them may embrace a utilitarian theory about the conditions under which ‘right’ applies, whereas the other may embrace a deontologist conception about this. We may also imagine cases where the antagonists embrace normative conceptions that differ more extensively, e.g. a theological view according to which these conditions consist in fulfilment of God’s will and a version of hedonism according to which they consist in satisfaction of purely sensuous pleasures. What is remarkable is that we conceive of people who have different—perhaps widely different—views of this kind to actually *disagree*. As regards non-moral terms, or at least as regards many such terms, we are less inclined towards such a conclusion. Consider a situation where two persons have quite different views about the conditions under which a certain non-moral term, ‘prime minister’ say, applies. For example, one holds that ‘prime minister’ applies to roughly the same individuals that we apply it to, whereas the other holds that the term applies to people who are, say, happy, rich and generous. A reasonable

---

<sup>21</sup> The view of persistent disagreements proposed in the present section—that they should be understood as expressions of conflicting normative conceptions—fits well with the idea suggested in the previous section, that moral disagreements, rather than being particularly widespread, are isolated phenomena which concern rather few but recurrent issues. Conflicting normative conceptions may often yield the same results—i.e. generally classify the same objects as having the same moral property—in most cases but come to contrary results in some others. It might then be argued that the cases where normative conceptions yield different results are precisely the rather few but recurrent cases about which there are persistent disagreements. It is in these cases the differences between various normative conceptions come to the surface.

<sup>22</sup> See e.g. Hare (1952), pp. 148–149. See also e.g. Blackburn (1984), p. 168; Schiffer (1990), pp. 609–610, and Tersman (2002), pp. 120–125. See also references to Hare below.

conclusion to draw in such a case would probably be that they do not disagree because they do not take the term to have the same meaning. That is, we seem to believe that, since they do not take the term to have the same meaning, they do not refer to the same property and consequently do not disagree.

Thus, what is remarkable about moral terms is that we are inclined to say that people who have (widely) different normative conceptions—and hence (widely) different views about the conditions under which a moral term applies—actually disagree, whereas we are less inclined towards a corresponding view concerning (many) non-moral terms. I will call this phenomenon *normative divergence*.

It may now be asked if the version of the argument from moral disagreement that appeals to the persistence of moral disagreements can be responded to in the same way as the version that appeals to their being widespread. It may seem so. Like the appeal to prevalence, the appeal to persistence involves a covert comparison with disagreements in other areas. And also in other areas there are presumably persistent disagreements. It might consequently be thought that in so far as we are not inclined to the conclusion that properties involved in these disagreements do not exist, we should not draw that conclusion as regards moral properties either. However, the phenomenon of normative divergence suggests that persistent moral disagreements cannot be dismissed in this way *even* if there is not more persistent disagreement in the moral area than in other areas.

The reason is that, due to normative divergence, it is a particular problem of explaining persistent moral disagreements in terms of an error in relation to a certain moral property. As was noted earlier, as regards (many) non-moral terms, we are inclined to say that people who hold (widely) different views about the conditions under which a term applies do not disagree *because they do not refer to the same property*. It may then be asked why we as regards *moral* terms are inclined to believe that people *do* disagree in spite of their having (widely) different views of this kind. Given our response in the non-moral case, it seems that the best explanation cannot involve the assumption that people involved in a disagreement about normative conceptions refer to the same moral property in relation to which some of them are in error. On the contrary, it seems that there has to be something else that provides the best explanation of their disagreement.

This reasoning finds support in a common way of thinking of the meaning of terms. We tend to assume that a term such as ‘prime minister’ applies to objects that we—at least if we are linguistically competent in respect of the term’s meaning—believe fulfil certain conditions in the form of having a certain set of properties. This set of properties is what we take to constitute the meaning of ‘prime minister’ and, hence, that to which the term refers. We also take this set of properties to be what makes up the property of being a prime minister. Consequently, if someone takes the term to apply to objects that have (widely) different properties than we take it to apply to, we assume that she does not refer to the property of being a prime minister, but to something else that she *calls* ‘prime minister’. And if we do not refer to the same property, we do not disagree. Applied to the moral case, this reasoning would mean that a moral term does not refer to a certain moral property error in respect of which we can explain disagreements concerning normative conceptions. Hence, it may be thought that such disagreements are best explained in some other way.

Moreover, it might be argued that the assumption that people who disagree about normative conceptions do not refer to the same property helps to explain why some moral disagreements are persistent. If people do not refer to the same property, it is only to be expected that they continue to disagree for the simple reason that they do not refer to a common subject matter in relation to which they can reach agreement.<sup>23</sup>

Thus, according to this reasoning, if feature F in premise (2) of the argument from moral disagreement is understood to consist in moral disagreements being persistent, there is a version of the argument that seems more persuasive than the one that appeals to such disagreements being widespread. As we have seen, typically persistent disagreements directly concern, or can be accounted for in terms of, disagreements belonging to type (iii): disagreements about normative conceptions. Such disagreements are characterised by the phenomenon of normative divergence. For the reason submitted above, it therefore seems reasonable to assume that the best explanations of such disagreements do not imply that there is an error in relation to a moral property. If this is correct, there are no moral properties, according to the argument from moral disagreement.

Non-cognitivists who employ the distinction between a primary and

---

<sup>23</sup> Cf. Tersman (2002), p. 123.



a secondary sense of moral terms may argue that they have the resources to provide a better explanation than one which involves the existence of moral properties. In fact, one of R. M. Hare's most recurrent arguments for his version of non-cognitivism rests on an argument to this effect.<sup>24</sup> Hare maintains that on the view that the meaning of a moral term is constituted by a set of non-moral properties, people who embrace different views about the conditions under which the term applies do not refer to the same property and hence do not disagree. He then appeals to the mentioned distinction to solve this problem. The primary sense of a moral term consists in a certain non-cognitive attitude, on Hare's view a prescription. The secondary sense of a moral term consists in certain non-moral properties. These properties constitute the person's moral criteria or the moral principles that determine towards which objects she has the non-cognitive attitude and hence to which object she applies the moral term. Formulated in terms of the vocabulary adopted above, they can be said to constitute her normative conception. Hare proposes that people may take the term to have the same primary sense, i.e. a non-cognitive attitude, but differ as regards the secondary sense they take it to have, i.e. differ in their normative conceptions. Accordingly, they may disagree in virtue of having conflicting non-cognitive attitudes while at the same time embrace different normative conceptions.

### *7. The First Reductionist Explanation of Persistent Moral Disagreement*

In the last section, we saw that a powerful version of the argument from moral disagreement is obtained if feature F in premise (2) is taken to consist in moral disagreements being persistent. Advocates of the existence of moral properties should then dispute premise (3) in this version of the argument. In order to do so, they should offer a plausible kind of explanation of persistent disagreements which involves the assumption that there is an error in relation to a moral property.

In the last section, we also saw that the basic reason why the persistence of moral disagreements gives rise to a significant version of the argument from moral disagreement is its connection to the phenomenon of normative divergence. This indicates that a satisfying explanation of

---

<sup>24</sup> See e.g. Hare (1981), pp. 68–71; Hare (1997), pp. 54–60, 68–70, 136–137; Hare (1999 (1993)), pp. 5–10, and Hare (1999 (1996)), pp. 83–85. See also e.g. Stevenson (1963 (1948)), pp. 1–3, 8–9, and Stevenson (1963 (1937)), pp. 26–29.

persistent disagreements has to comprise an account of this phenomenon. On the assumption that there are moral properties, this means that such an explanation has to comprise an account of how people who embrace (widely) different normative conceptions can refer to the same moral property and so how they can disagree in a way that is explainable in terms of the mentioned kind of error.

For reductionism primarily two types of explanations of moral disagreements are available. As we noted earlier, reductionism can be understood to state that the meaning or reference of a moral term ‘M’ is constituted by a non-moral property G, where G is identical with a moral property M.<sup>25</sup> This implies that whether objects have M depends on two things: what non-moral property constitutes the meaning or reference of ‘M’ and whether the objects have that property. Thus, if people are involved in any of the kind of disagreements mentioned in section 3, a reductionist explanation should primarily be either that they are in error about the meaning or reference of ‘M’, or that they are in error about some non-moral fact, i.e. some fact pertaining to the mentioned non-moral property.<sup>26</sup>

The two main forms of reductionism provide different explanations of the first kind, whereas the second kind of explanation is common to both. In this section, I will examine the first kind of explanation and in the next the second kind of explanation.

### *Analytic reductionism*

According to analytic reductionism, the meaning of a moral term ‘M’ is constituted by a non-moral property G. The distinctive explanation analytic reductionism offers for why people morally disagree is provided in terms of an error concerning the meaning of the term thus understood.

On analytic reductionism, the explanation that some people who are

---

<sup>25</sup> I use the phrase ‘meaning or reference’ to indicate that reductionism exists in basically two versions: analytic and synthetic reductionism.

<sup>26</sup> There are other explanations that reductionists may employ to account for persistent moral disagreements. For example, they may propose that the error in question is due to a failure to draw the correct conclusion from the relevant premises. I doubt, however, that these explanations are of the significance that they would have any important implications for my arguments against reductionism. In what follows, I will argue that the two main types of explanations reductionists can employ fail. Even if there are other kinds of explanations that they can appeal to, I doubt that they would be sufficient to show that reductionism succeeds to account for persistent moral disagreements.

involved in moral disagreements are mistaken about the meaning of a moral term indicates that they lack linguistic competence in respect of the term in the sense of not knowing what constitutes its meaning. This suggestion generates two problems.

Firstly, concerning most moral disagreements, it is implausible to maintain that people are not linguistically competent in the indicated way. Most noteworthy in the present context, it is untenable to understand those who hold erroneous normative conceptions as not knowing what the term in question means. For example, on the assumption that utilitarianism is the correct normative conception concerning rightness, we believe that those who oppose this view, deontologists for instance, are mistaken, but not that they do not know what 'right' means.

Secondly, the explanation under discussion suggests that moral disagreements are merely apparent. As we have seen, people who are mistaken about the meaning of the moral term at issue are not linguistically competent with respect to the term in the sense that they do not know what non-moral property constitutes its meaning. The proposed explanation then suggests that they take the term to refer to something else than it actually does. This, in turn, suggests that they and the people with whom they debate do not refer to the same property and hence do not disagree, but just talk past each other. This is unreasonable since we certainly conceive of the participants in such a debate as actually disagreeing. In particular, in relation to debates about normative conceptions, it is unreasonable to take, for instance, utilitarians and deontologists not to disagree, but just to talk past each other.

The last point is particularly important in relation to the phenomenon of normative divergence. As regards (many) non-moral terms, we are inclined to accept that people who have (widely) different views about the conditions under which such a term applies do not disagree. However, as we saw in the last section, this is not a conclusion towards which we are inclined when it comes to moral terms. The explanation analytic reductionism offers is therefore especially untenable with regard to such terms. Consequently, this explanation is unable to account for normative divergence.

According to David O. Brink, it is possible to disagree with people who embrace a moral view that on analytic reductionism would be

expression of the meaning of a moral term. He writes:

Suppose most speakers associate features X, Y, and Z with general term 'G'. It ought to be possible for a heretical inquirer to express disagreement with the prevailing view. A speaker ought to be able to say that the very thing that most speakers use 'G' to refer to is not X, Y and Z, but rather is A, B and C (where 'X, Y and Z' and 'A, B, and C' have different extensions). However, this is ruled out by the traditional descriptive theory, for on this view the meaning and reference of 'G' is given by the description—X, Y, and Z—that is conventionally associated with 'G'. The heretic's claim would thus be analytically false. But certainly not all heretical claims are false—much less analytically false—as the progressive nature of various inquiries shows us.<sup>27</sup>

Brink's argument can be understood in the following way. According to analytic reductionism as Brink understands it, the meaning of a moral term consists, roughly put, in the set of non-moral properties that competent speakers conventionally associate with the term. In Brink's view, it is possible for a person to disagree with people who embrace the prevailing moral conception and consequently to reject the conception expressed in the assumed meaning of such a term. That this is possible is shown by the fact that, through history, what was once the received moral opinion about various moral issues has later been reappraised and considered to be mistaken. However, this is impossible on analytic reductionism as Brink understands it, since a person who rejects such a view would have to take the term to refer to quite another set of non-moral properties than the one which constitutes its meaning. Hence, Brink claims, analytic reductionism is mistaken. In sections **10** and **11**, we will see that, according to realism, it is possible to disagree about conceptions even if they express the received moral opinion.

### *Synthetic reductionism*

According to synthetic reductionism, the reference of a moral term 'M' consists in a non-moral property G. The distinctive explanation synthetic reductionism offers for why people morally disagree is provided in terms of an error as regards the reference of the term thus understood.

This explanation does not suffer from the two main difficulties with the explanation analytic reductionism offers of moral disagreements. First, according to synthetic reductionism people who are mistaken about the reference of a moral term need not lack linguistic competence with regard

---

<sup>27</sup> Brink (2001), p. 159.

to it. On this view, it is not part of this competence to know what non-moral property constitutes its reference. This is something that is found in an investigation, the results of which linguistically competent persons do not necessarily have access to. Second, the explanation under discussion does not imply that moral disagreements are merely apparent because people refer differently with the moral term in question. As just mentioned, on this view it is not part of being linguistically competent with respect to a moral term to know what non-moral property constitutes its reference. This means that people may refer to this non-moral property with a moral term without necessarily knowing which this property is and even if they think that it consists in something quite else than in actually does. This leaves open the possibility that people may refer to the same non-moral property as those with whom they disagree while erroneously believing that the moral property consists in something else. Hence, the explanation under consideration allows for genuine moral disagreements.

Moreover, as the second comment suggests, synthetic reductionism seems able to account for the phenomenon of normative divergence.<sup>28</sup> According to this view, in contrast to analytic reductionism, people who embrace (widely) different normative conceptions, and hence have (widely) different views about the conditions under which a moral term applies, may refer to the same moral property. As suggested above, this is so since they may refer to the same non-moral property with the moral term in spite of the fact that some of them are mistaken about its reference and hence are mistaken about what conditions should be fulfilled for the term to apply. It might therefore be argued that on synthetic reductionism disagreements about normative conceptions are explainable in terms of an error in relation to a moral property.

However, I think it can be argued that synthetic reductionism has certain shortcomings.

Firstly, synthetic reductionism opposes our notion of what people who are involved in disagreements about normative conceptions refer to with moral terms. In chapter 3, I argued that on synthetic reductionism people refer with a moral term to a non-moral property that is specified by a

---

<sup>28</sup> This is often thought to be one of the main advantages of synthetic reductionism; see e.g. Adams (1981 (1979)), p. 112; Boyd (1988), pp. 209–210; Brink (2001), pp. 162–163; Goldman (1988), p. 153; Merli (2002), p. 208, and Sayre-McCord (1997a), pp. 269, 276, 281. Cf. Loeb (1998), pp. 294–295.

particular normative theory. This view proves to be problematic in the cases where the antagonists in a moral disagreement embrace specific normative conceptions. To see this, consider the following example. Suppose that, as regards rightness, the normative theory in question is utilitarianism so that 'right' refers to maximising happiness. Suppose further that a convinced utilitarian debate with a person who is equally convinced that utilitarianism is mistaken and that quite another normative conception is correct, e.g. a dedicated deontologist. According to synthetic reductionism, the deontologist then refers to maximising happiness when she uses 'right'. Moreover, she does so without knowing it and in spite of her believing that rightness consists in something quite else. Consequences like this oppose, I think, our view of what people involved in disagreements over normative conceptions refer to with moral terms. Accordingly, I do not think we understand a person who is a sworn opponent to utilitarianism and convinced that another normative conception is correct to be referring to maximising happiness with 'right'. In so far as we believe that she refers to a non-moral property that is specified by a particular normative theory at all, we take her to refer to quite *another* non-moral property, namely the one specified by the normative theory she embraces, in case there is such a theory. The deontologist would quite plausibly agree; in particular, she would strongly insist that she does not refer in a way that presumes the truth of utilitarianism.

Secondly, synthetic reductionism construes the contents of certain moral sentences accepted by participants in disagreements about normative conceptions in a way that opposes our notion of what the contents of these sentences consist in. This difficulty also has its origin in the view that people refer with a moral term to a non-moral property which is specified by a particular normative theory. To see this difficulty, imagine once again that a utilitarian and a deontologist are engaged in a debate about their normative conceptions. Suppose also that the normative theory that specifies the reference of 'right' is utilitarianism so that the term refers to maximising happiness. Let us first consider a kind of case that Ernest Sosa calls attention to.<sup>29</sup> Suppose the deontologist utters the following sentence: 'Only actions that comply with the deontological conception are right'. According to synthetic reductionism, she has then produced a sentence with the following

---

<sup>29</sup> Sosa (1997), p. 310.

content: ‘Only actions that comply with the deontological conception maximise happiness’. As Sosa remarks, ‘This is hard to believe’. However, this is not the end of the matter. Suppose the utilitarian says: ‘Actions that maximise happiness are right’. According to the view under consideration, she has then produced a tautology: ‘Actions that maximise happiness are such that they maximise happiness’. Moreover, suppose the deontologist utters the following sentence: ‘Actions that are right don’t maximise happiness’. According to synthetic reductionism, she has then produced a contradiction: ‘Actions that maximise happiness don’t maximise happiness’. Also this is hard to believe. As noted several times, on synthetic reductionism people may refer to a non-moral property with a moral term without necessarily knowing which this property is. Hence, the participants in the disagreement need not, and presumably do not, understand themselves to utter sentences that have these contents. Yet, I think synthetic reductionism construes the contents of these sentences in a way that opposes our notion of what these sentences plausibly can be taken to say.<sup>30</sup>

These two arguments have implications for the prospect of synthetic reductionism to account for normative divergence. Normative divergence concerns disagreements about normative conceptions. Since the two arguments indicate that synthetic reductionism misconstrues disagreements about normative conceptions, there is reason to believe that it fails to account for normative divergence.

Advocates of synthetic reductionism may argue that they can counter these objections by providing an explanation of why it is reasonable to understand people to be referring to a non-moral property with a moral term in the way synthetic reductionism states. As we saw in chapter 3, synthetic reductionism claims that we refer to a certain non-moral property with a moral term in so far as our use of the term is causally regulated by that non-moral property. Return to the example of a disagreement

---

<sup>30</sup> It should be noted that according to the view of the reference of natural kind terms mentioned in chapter 3, a person who uses a natural kind term, e.g. ‘water’, refers to a natural kind, e.g. H<sub>2</sub>O, without necessarily knowing it and even if she believes that what she refers to consists in something else. Hence, this view has implications similar to that of synthetic reductionism. I will not discuss this issue here, but two comments are worth making. First, I think implications like those considered above are intuitively more problematic as regards moral terms than as regards natural kind terms. Second, it might be argued that there are differences between natural kind terms and moral terms that explain this. According to one such explanation, moral terms have quite other functions than natural kind terms.

concerning normative conceptions above. There are considerable differences between the normative conceptions defended by the utilitarian and the deontologist. However, in spite of these differences, it might be maintained that their use of the term is causally regulated by the same non-moral property, in the example above that of maximising happiness. More precisely, it might be argued that maximising happiness is causally responsible for certain reference-fixing characteristics that both of them utilise to apply 'right' and so that both their use is causally regulated by this non-moral property. This explanation, synthetic reductionists may contend, makes it reasonable to claim that people who embrace different normative conceptions do refer to the same non-moral property, despite the two objections mentioned above.

However, rather than providing a plausible response to the two objections, I think this argument suggests that synthetic reductionism is vulnerable to an additional objection. There is reason to believe that people may disagree about their normative conceptions even if these conceptions are so different that it is implausible to assume that their use of the moral term in question is causally regulated by the same non-moral property.<sup>31</sup> This is problematic for synthetic reductionists because they are committed to the view that people whose use of a moral term is not causally regulated by the same non-moral property—for example because they embrace widely different normative conceptions—do not refer to the same property and hence do not disagree. This indicates that synthetic reductionism is unable to account for how people with (widely) different normative conceptions can disagree. It reinforces thus the contention suggested above: that synthetic reductionism is incapable of accounting for normative divergence.

To illustrate the mentioned kind of situation, consider a case where people debate about normative conceptions that differ more thoroughly than the normative conceptions embraced by the utilitarian and the

---

<sup>31</sup> In this regard moral terms might be contrasted with natural kind terms like 'water'. I think we might be prepared to accept that if people's beliefs differ widely in the relevant respect, their use of the term 'water' is not causally regulated by the same stuff (H<sub>2</sub>O) and that they therefore do not refer to the same thing. In the moral case, we might be inclined to say that if people's normative conceptions differ widely, their use of the term 'right' is not causally regulated by the same non-moral property. However, we do not take this to mean that they do not refer to the same property and that they cannot disagree as regards rightness.



deontologist.<sup>32</sup> Imagine, for example, that a person advocates a theological normative conception according to which the rightness of actions consists in their fulfilling God's will, where this is understood in terms of the Old Testament. Suppose she debates her moral outlook with a person who advocates a version of hedonism according to which the rightness of actions consists in their bringing about purely sensuous pleasures. Assume further that they follow their respective normative conceptions reliably in the sense that they apply 'right' in accordance with their respective view. Considering the difference between their normative conceptions, they presumably apply 'right' rather differently. In such a case, it does not seem plausible to assume that their use of 'right' is causally regulated by the same non-moral property. In particular, it does not seem plausible to assume that a particular non-moral property—such as maximising happiness—is causally responsible for certain reference-fixing characteristics which both of them utilise when they apply the term. On synthetic reductionism, this would mean that they do not refer to the same non-moral property with the term and therefore do not disagree. Yet, I think we are inclined to say that they do disagree and consequently refer to the same property with 'right'.

Advocates of synthetic reductionism may reply to this argument by maintaining that there are explanations of how people can have their use of a moral term causally regulated by the same non-moral property in spite of their embracing widely different normative conceptions. In particular, they might argue that people's beliefs about the relevant non-moral facts can differ in such a way that the mentioned kind of situation occurs.

It should be admitted that it is difficult to outright demonstrate that such explanations are not available. However, given that people's normative conceptions may differ extensively and that they are faithful to these views when they apply the moral term in question, it seems questionable whether such differences generally are explainable in the suggested way. Moreover,

---

<sup>32</sup> This argument is similar to the one directed against synthetic reductionism in chapter 3 and to the argument Horgan and Timmons direct against this view; see e.g. Horgan and Timmons (1992b), pp. 240–248. However, the arguments are distinct. The former argument is, briefly put, that people whose use of a moral term is causally regulated by different non-moral properties may disagree, and they refer consequently to the same property. My argument here is, briefly put, that the difference between people's normative conceptions might be so extensive that it is implausible to assume that their use of the moral term in question is causally regulated by the same non-moral property. However, in spite of this, they may disagree, and they refer consequently to the same property.

in the next section I will argue that people who are involved in moral disagreements, particularly those involved in disagreements about normative conceptions, may take a moral term to refer to the same property and yet not differ as regards any relevant non-moral beliefs.

### 8. *The Second Reductionist Explanation of Persistent Moral Disagreement*

As we have seen, for reductionists there are two principal ways to explain moral disagreements: an error as regards meaning or reference of a moral term and an error as regards non-moral facts. In the last section, I argued that the first kind of reductionist explanation fails to provide the required explanation of persistent disagreements. In this section, I will investigate whether the second kind of reductionist explanation succeeds in doing so.<sup>33</sup>

Let us first notice an implication of this kind of explanation. Suppose two persons disagree in any of the ways mentioned in section 3. Suppose further that neither of them is mistaken as regards the meaning or reference of a moral term. Reductionists are then likely to suggest that the disagreement is explainable in terms of an error as regards some non-moral fact. This explanation implies that people who disagree differ concerning at least one non-moral belief pertaining to such a fact.

It should be admitted that an error as regards non-moral facts presumably provides a plausible explanation of a number of moral disagreements. We may assume that it explains many disagreements of type (i); for instance, the fact that people disagree as to whether a particular action is right can be explained by some of them having erroneous non-moral beliefs about it. The reason why people entertain erroneous beliefs of this kind might for example be that they are influenced by self-deception and egoistic motives. More interestingly, errors as regards non-moral facts also seem suitable to explain a number of widespread disagreements of type (ii); for instance, the fact that people in different cultures disagree as regards the rightness of certain types of actions might be explained by some of them having erroneous non-moral beliefs about these actions. People might for example entertain erroneous non-moral beliefs about the relevant consequences of a certain kind of actions because knowledge about these consequences is very hard to acquire or because they are influenced by

---

<sup>33</sup> This kind of explanation is proposed by e.g. Brink (1989), pp. 202–203; Ross (1939), pp. 16–17, and Shafer-Landau (2003), pp. 218–219. This kind of explanation is criticised by e.g. Hare (1997), pp. 68–69, 73–78, and Loeb (1998), pp. 283–284, 290.

certain culturally transmitted prejudices.

However, there is one significant difficulty with this kind of explanation. The basic reason is that it seems that people who are involved in moral disagreements can take a moral term to have the same meaning or reference and yet not differ concerning any relevant non-moral beliefs.

Let us first notice, as Charles L. Stevenson in particular has drawn attention to, that the mentioned kind of situation is conceivable.<sup>34</sup> The following example can be used to illustrate this. Suppose that two persons have disputed for quite a while as to whether a certain action, or a certain kind of actions, is right without having reached agreement. They both take 'right' to have the same meaning or reference, something that is confirmed by the fact that both of them are considered as linguistically competent in respect of them term. Since they believe that they take 'right' to have the same meaning or reference, they assume that the explanation of their disagreement is to be found in their differing as regards some relevant non-moral beliefs. They then ask themselves if they have different beliefs about, for example, the consequences of the type of actions under dispute or the intentions with which such actions typically are performed. But they do not find any non-moral belief that explains why they disagree. As a matter of fact, there is no such belief; they agree about all relevant non-moral beliefs.

It seems reasonable to suppose that situations as the one described are not merely conceivable, but actually occur at times. Sometimes we seem to be involved in moral disputes where we take the moral term in question as having the same meaning or reference, and hence genuinely disagree, but yet cannot find any difference in relevant non-moral beliefs. It is notable that in such situations we are inclined to say that we disagree because we embrace fundamentally different views about moral matters. Formulated in the terms employed here, we disagree because we embrace different normative conceptions. This may be taken to suggest that disagreements about normative conceptions cannot be explained merely in terms of differing non-moral beliefs.

It is notable that as regards non-moral terms, or at least as regards many non-moral terms, the mentioned kind of situation does not seem to

---

<sup>34</sup> See Stevenson (1944), esp. chap. 1; Stevenson (1963 (1938)), pp. 51–54, and Stevenson (1963 (1948)), pp. 5, 7–9. This view is also suggested by Hare; see previous references. See also e.g. Darwall (2003), p. 11; Loeb (1998), p. 290; Tersman (2002), p. 32, and Wedgwood (2001), pp. 27–29.

be conceivable, much less occur. Suppose, for example, that two persons disagree as to whether a certain individual is a prime minister. On the assumption that they both take the term 'prime minister' to have the same meaning or reference, it seems that they have to differ as regards some relevant 'non-prime minister belief'. It seems difficult to come to think of a situation where people disagree whether a certain individual is a prime minister, take the term to have the same meaning or reference, and yet do not differ in the mentioned respect.

As Stevenson suggests, the observed phenomenon may be thought to give support to non-cognitivism. The reason is that people on this view may hold the same non-moral beliefs and yet disagree morally in virtue of differing with regard to the non-cognitive attitude that constitutes the (primary) meaning of the moral term in question.<sup>35</sup>

Above it was suggested that the mentioned type of phenomena is conceivable, and may occur, in relation to disagreements as to the moral property of a particular object or a certain kind of objects, i.e. disagreements of type (i) and (ii). However, it is presumably even more likely to occur in relation to disagreements about normative conceptions, i.e. disagreements of type (iii). This is important in the present context, since persistent disagreements typically concern, or are explainable in terms of, disagreements about normative conceptions. It seems especially implausible to characterise disagreements about normative conceptions in a way that makes them explainable merely in terms of differing non-moral beliefs. The main reason is that disagreements about such views, rather than being concerned with purely non-moral facts, seem fundamentally moral in nature. As D. Loeb remarks, 'consider the debate between Kantians and utilitarians. It seems very unlikely that the continued existence of this debate hinges upon disagreement over the non-moral facts.'<sup>36</sup>

This contention finds support if we consider some of the arguments presented in debates about normative conceptions. As an illustration, reflect on the well-known arguments put forward against utilitarianism, e.g. that it threatens personal integrity and special relationships like love and friendship or that it is too demanding. It would be far-fetched to consider these arguments as intended to point out that some purely non-moral facts are or

---

<sup>35</sup> See references to Stevenson above.

<sup>36</sup> Loeb (1998), p. 290.

are not the case. Rather, these arguments seem typically to concern *moral* notions in the sense that they are meant to call attention to which non-moral properties, in some sense, contribute to, or fail to contribute to, actions being right. Put somewhat more exactly, they seem primarily intended to draw attention to either of two types of cases. Either that a certain non-moral property (e.g. keeping a promise) contributes to actions being right, but that this fact is not accommodated by utilitarianism. Or that a certain non-moral property does not, at least not always, contribute to actions being right, but that utilitarianism implies so; that is, it might be argued that what contributes to actions being right is not that they maximise happiness, at least not always.

### 9. *A Note on Relativism*

In the preceding two sections, I have argued that reductionists are not successful in rejecting the pertinent version of the argument from moral disagreement by opposing premise (3) of the argument. At this point, reductionist may want to take a step back and question something that is tacitly presumed in premise (1), namely that there are any moral disagreements of the relevant kind in the first place. This response to the argument may be proposed by reductionists who advocate relativism.<sup>37</sup>

We may start with considering, in bare outline, what the operative feature of the relevant type of relativism amounts to. To illustrate, suppose that a person asserts a moral sentence in which rightness is ascribed to an action. According to the type of relativism at issue, a person who appears to deny the sentence may not really be doing so. The reason is that morality is relative, with the consequence that she uses ‘right’ in such a way that it has a different meaning or reference than the term has when it is used by the person who asserts the sentence. To put it very roughly, one person uses the term in such a way that its meaning or reference consists in one non-moral property, whereas the other uses it in such a way that its meaning or reference consists in another non-moral property.<sup>38</sup> Accordingly, morality is relative in a way that allows that they do not actually disagree as to the

---

<sup>37</sup> For this line of reasoning, see e.g. Harman in Harman and Thomson (1996), pp. 8–19, and Wong (1984), esp. chap. 9–11.

<sup>38</sup> However, this should not be understood to imply that they are aware that their usage of ‘right’ differs in this way.

rightness of the action, although it might appear that they do so disagree.<sup>39</sup>

This characterisation of the operative feature of the relevant kind of relativism needs undoubtedly to be refined in a number of ways. Especially, it should be made explicit what the relativity of morality amounts to. This can be done in a variety of ways. According to Gilbert Harman, morality is relative because of the existence of various ‘moral frameworks’. More precisely, ‘a judgment of the form, *it would be morally wrong of P to D*, has to be understood as elliptical for a judgment of the form, *in relation to moral framework M, it would be morally wrong of P to D*. Similarly for other moral judgments.’<sup>40</sup> According to the version of relativism that is relevant here, ‘framework M’ can be taken to consist in a normative conception which states that the moral property in question, e.g. rightness, is constituted by a certain non-moral property. It may then be assumed that a moral term such as ‘right’ varies in meaning or reference depending on which the normative conception in question is. Moreover, the present version of relativism says that there are various normative conceptions, which claim that the moral property in question is constituted by different non-moral properties, and that these conceptions are equally correct.

How the relevant type of relativism makes use of the operative feature mentioned above in order to reject the argument from moral disagreement depends partly on which kind of moral disagreement that is appealed to in the version of the argument at issue. On the version of the argument which concerns us here, this is persistent disagreements. Such disagreements are typically expressions of different normative conceptions, as we have seen earlier. As just indicated, relativists may then claim that people who embrace different normative conceptions use ‘right’ in such a way that its meaning or reference consists in different non-moral properties. It follows that they do not disagree. It can then be argued that, contrary to what premise (1) implicitly suggests, there are no relevant moral disagreements, i.e. no disagreements of the sort that is appealed to in the present version of the argument from moral disagreement. Thus, it might be

---

<sup>39</sup> Of course, they may disagree in other respects; for example, they may disagree in the sense of having diverging attitudes.

<sup>40</sup> Harman in Harman and Thomson (1996), p. 4. David Wong proposes similar analyses for ‘morally ought’ and ‘morally good’: Wong (1984), chap. 4–6. Cf. Garcia (1988), pp. 269–271; Lyons (1982), pp. 211–213, and Sayre-McCord (1991), pp. 161–163. Needless to say, there are other versions of relativism than the one considered here; for example, relativism need perhaps not be reductionist.

maintained that there are no relevant disagreements that are to be best explained in terms of an error with respect to a moral property. Hence, the argument cannot be used to show that there are no moral properties.

Relativism undoubtedly requires much more attention than I am able to give it here. However, some simple standard arguments indicate that reductionism should not make use of this view to respond to the argument from moral disagreement.<sup>41</sup>

Firstly, it runs counter to one of our fundamental convictions about morality, namely that there are moral disagreements of the mentioned kind. Suppose two persons debate whether a certain type of actions, e.g. abortion, is right. Surely, we conceive of them as disagreeing and not only talking past each other because they embracing different normative conceptions.<sup>42</sup> In the like manner, we conceive of a person who asserts a sentence to the effect that abortion is right to be agreeing with a person who accepts the sentence, even if they embrace different normative conceptions. The same reasoning applies in the cases where people debate normative conceptions as such. For instance, we take the utilitarian and the deontologist to be disagreeing, not just talking past each other.

In this connection, it should also be pointed out that relativism is incompatible with normative divergence, since it implies that people involved in debates about normative conceptions do not disagree.

Moreover, relativism suggests implausibly that debates about moral matters are pointless. On this view, it would be pointless for a person to debate whether a certain type of actions is right with someone who

---

<sup>41</sup> It might perhaps be doubted whether relativism is an instance of reductionism. On a relativist view of a moral property such as rightness, there is more than one non-moral property that constitutes the meaning or reference of 'right'. There is thus not one moral property, rightness, which is identical with a particular non-moral property. Rather, for each meaning or reference of 'right' there is a non-moral property. There is thus one property, *right*<sub>1</sub>, which is identical with a certain non-moral property, a further property, *right*<sub>2</sub>, which is identical with another non-moral property, and so forth. However, I will not dwell on this issue as it does not seem particularly important to the present discussion. In this context, it can also be noted that just as there might be a relativist version of reductionism, there might perhaps be a relativist version of realism. On this view, there are several moral properties of a certain kind, e.g. *right*<sub>1</sub>, *right*<sub>2</sub>, and so forth, where each such moral property is irreducible to a non-moral property. Hence, it is not the case that each such moral property is identical with a certain non-moral property. However, for each of these moral properties, there might be analogues to the realist formula which state that different sets of non-moral properties make actions *right*<sub>1</sub>, *right*<sub>2</sub>, etc.

<sup>42</sup> Of course, this is a common objection against relativism; see e.g. Lyons (1982), p. 222, and Sturgeon (1994), pp. 97, 99–100.

advocates another normative conception for the simple reason that they would not disagree but only talk past each other. For the same reason, it would be pointless for people who embrace different normative conceptions to argue about these conceptions as such. However, we find moral debates meaningful to pursue even when we discuss with people who hold quite different normative conceptions than we do. For example, a utilitarian finds it presumably meaningful to debate whether abortion is right with someone who judges abortion differently because she holds another normative conception; indeed, she may even find it more to the point to debate with someone who holds a different normative conception. Likewise, advocates of different normative conceptions find it meaningful to debate normative conceptions as such.

In reply to these objections, relativists may point to two considerations that they believe talk in favour of their view.

First, they may argue that they are able to explain what is troubling with persistent disagreements.<sup>43</sup> Relativists may explain how it can be that people debate certain moral issues extensively without reaching agreement by suggesting that they talk about different things; if people talk about different things, it is not very surprising that they do not agree, as they do not debate some common subject matter about which they can reach agreement.

In response to this argument, it should be pointed out that even if relativism can explain why people involved in persistent disagreements do not reach agreement, it is unable to explain why they *continue* to debate. Assuming relativism, if people have debated an issue extensively without having reached agreement, it seems that they should have come to the conclusion that they talk about different things and accordingly brought the debate to an end.

Moreover, advocates of relativism may argue that this view offers the only alternative for those who want to uphold the existence of moral properties. As we have seen in the preceding two sections, non-relativist reductionism fails to respond to the version of argument from moral disagreement at issue. Thus, if we adhere to this view we have to accept the conclusion of the argument: that there are no moral properties. Relativists avoid the argument by arguing that there are no disagreements of the

---

<sup>43</sup> Cf. Harman in Harman and Thomson (1996), p. 11.



relevant kind. They may admit that with relativism follows some undesired consequences, but maintain that these have to be accepted to save the existence of moral properties.

This argument rests on the presumption that there is no other way to reject the argument from moral disagreement than those offered by non-relativist reductionism and relativism. Against this presupposition, I will in the next three sections argue that realism provides a way of disputing the argument from moral disagreement under consideration which does not have the drawbacks of either of these views.

### **10.** *A Realist Account of Normative Divergence*

In section 6, we saw that if feature F in premise (2) of the argument from moral disagreement is assumed to consist in moral disagreements being persistent, a significant version of this argument is obtained. In order to reject it, advocates of the existence of moral properties should dispute premise (3) of the argument thus understood by offering a plausible explanation of persistent disagreements in terms of an error in relation to a moral property. In sections 7 and 8, I argued that reductionism fails to provide such an explanation. In the previous section, I argued that reductionism should not take support on relativism to evade the argument. In this and the two subsequent sections, I will argue that realism is able to provide the required explanation and hence is capable of rejecting premise (3).

We noticed in section 6 that the main reason why the persistence of moral disagreements can be appealed to in the argument from moral disagreement is its close connection to normative divergence. For this reason an explanation of persistent disagreements has to comprise an account of this phenomenon. We also noticed that the failure of reductionism to account for normative divergence is one of the main reasons why it is unable to explain persistent disagreements. I start therefore in this section with a realist account of normative divergence and continue in the two next sections with a realist explanation of persistent disagreements.

Normative divergence, we may recall, is the phenomenon that people who hold (widely) different normative conceptions—and hence have (widely) different views about the conditions under which a moral term applies—nonetheless may disagree. By contrast, as regards (many) non-

moral terms, we regard people who hold (widely) different views about the conditions under which such a term applies not to be disagreeing because they take the term in case to have different meanings and, hence, do not refer to the same property. It might therefore be argued that normative divergence means that persistent disagreements needs to be explained in some other way than in terms of an error in relation to a moral property.

On the realist account of normative divergence, people who hold different normative conceptions may take the moral term in question to have the same meaning and may accordingly refer to the same property. Despite what the comparison with non-moral terms suggests, their disagreements can thus be explained in terms of an error in relation to a moral property. The suggestion of how this is possible is, roughly put, that the conditions under which people employ a moral term so as to have the same meaning and, hence, refer to the same property, are, in a certain respect, less specific as compared with (many) non-moral terms.

When it comes to non-moral terms, the meaning of a term might be stated by necessary and sufficient conditions that yield an analytically necessary biconditional. This suggests that people who have (widely) diverse views about these conditions do not refer to the same property with the term. Admittedly, the conditions that constitute the meaning of a non-moral term need not be so strict so as to yield an analytically necessary biconditional. Yet, concerning (many) non-moral terms, they are so strict that it is not possible for people to have different views—at least not widely different views—about the conditions under which such a term applies and still refer to the same property. Hence, it is not possible for people who hold (widely) different views of this kind to disagree.

However, the conditions under which people employ a moral term to have the same meaning, and hence refer to the same property, are not specified in that way. In particular, there are reasons to believe that, contrary to what analytic reductionism implies, no analytically necessary biconditional involving a moral property and a non-moral property holds. Instead the relevant aspect of the meaning of moral terms is captured in the realist formula.

We may illustrate how the realist account of normative divergence works as regards the meaning of a particular moral term, ‘right’. In relation to this term, the realist formula says, roughly put, the following: (i) It is

analytically necessary that if an action is right, there is a set of non-moral properties G such that **(A)** the action has G, and **(B)** it is synthetically necessary that whatever action has G is right. **(ii)** Rightness is not identical with any non-moral property.

As mentioned in the previous chapter, I think it is reasonable to suggest that the set of non-moral properties mentioned in the realist formula is of the kind specified by normative conceptions. Thus, on the realist formula, when people assert that an action is right, they can be understood to claim that it has a set of non-moral properties—i.e. a set of non-moral properties of the kind involved in a normative conception—such that it is synthetically necessary that whatever action has that set of non-moral properties is right. In the last chapter, it was also implied that on the realist formula, there is not one particular set of non-moral properties that actions must have whenever they are right. Rather, on this view people who assert that an action is right can be understood to claim that the action has *some* set of non-moral properties—some set of non-moral properties of the kind involved in a normative conception—such that it is synthetically necessary that whatever action has that set of non-moral properties is right.

Thus, that people use ‘right’ in accordance with the realist formula does thus not imply that they refer to a particular set of non-moral properties that is involved in a normative conception. On the contrary, the formula allows that people who advocate different—even widely different—normative conceptions nevertheless may refer to the same property with the term. It is accordingly possible for them to disagree in respect of rightness in spite of their embracing (widely) different normative conceptions.

Above I suggested that the set of non-moral properties mentioned in the realist formula is of the kind specified by normative conceptions. According to this view, people who disagree about normative conceptions consequently disagree about what such a set of non-moral properties consists in. In the preceding chapter, I mentioned that the realist formula can be understood to say that a set of non-moral properties *makes* objects have a moral property, such as rightness. On this account, advocates of different normative conceptions can accordingly be said to disagree about what makes actions right. This suggestion can be specified further. It might be proposed that implications like **(B)** in the realist formula are expressions of normative conceptions. That is, advocates of different normative

conceptions can be understood to have differing views of what the set of non-moral properties referred to in such an implication consists in and consequently whether such an implication holds.<sup>44</sup>

On the realist account, a disagreement between two persons about normative conceptions with regard to rightness can be understood to involve two aspects. First, both of them refer with ‘right’ in accordance with the realist formula. Roughly put, in asserting that an action is right, they claim that it has a set of non-moral properties—some set of non-moral properties of the kind involved in a normative conception—such that it is synthetically necessary that if an action has that set of non-moral properties, it is right. In virtue of this, they refer to the same property, rightness, and can consequently disagree. Second, they disagree as to whether a certain set of non-moral properties makes actions right. More exactly, they disagree as to whether a synthetically necessary implication of the kind **(B)** holds. In virtue of this, they disagree about normative conceptions.

### *11. A Realist Explanation of Persistent Moral Disagreement*

In the previous section, I argued that realism is able to account for normative divergence. In the present section, we can continue with considering how realism can explain persistent moral disagreements.

Let us start by briefly recalling the two primary kinds of explanations reductionists offer of moral disagreements. According to reductionism, the meaning or reference of a moral term ‘M’ is constituted by a non-moral property G. This means that whether objects have a moral property M depends on two things: what non-moral property constitutes the meaning or reference of ‘M’ and whether objects have that non-moral property. Reductionists may then suggest that people disagree because of an error as regards the meaning or reference of the term or because of an error as regards some non-moral fact pertaining to the mentioned non-moral property. It might be observed that only the first of these considerations concerns what constitutes the moral property.

According to the realist formula, the meaning or reference of a moral term does not consist in a particular non-moral property. This means, as we shall see, that realism is able to explain persistent moral disagreements in a

---

<sup>44</sup> This account should not be taken to imply that they consciously think of the disagreement in terms of such an implication.

manner that directly concerns what constitutes a moral property in another way than reductionism.

Let us now consider how realism can account for persistent moral disagreements. Because of the tight connection between such disagreements and normative divergence, what was said in the last section has already given a hint of this explanation. I will use ‘right’ to illustrate this explanation, but I think it applies to other moral terms as well.

According to realism, people involved in a persistent moral disagreement in relation to rightness take ‘right’ to have the meaning and reference as stated in the realist formula. Hence, the error that explains the disagreement does not consist in an error concerning the meaning or reference of the term. Neither does it consist in an error concerning non-moral facts, so the explanation of the disagreement is not to be found there either. Rather, what explains their disagreement is an error concerning what makes actions right, where ‘make’ is understood in accordance with the realist formula. More precisely, it is an error as regards the kind of implication (**B**) in the realist formula, i.e. an error as to whether it is synthetically necessary that if actions have a certain set of non-moral properties, it is right.

We may now see that the realist explanation does not have the difficulties of the reductionist explanations. Let us first see why this is so in relation to *the first kind of explanation* reductionism employs.

Consider first this kind of explanation when employed by *analytic reductionism*. Analytic reductionism has, as we saw in section 7, mainly two problems in relation to this kind of explanation. First, it suggests that some people involved in debates concerning rightness are linguistically incompetent in respect of the meaning of ‘right’. Second, it suggests that people involved in such debates do not really disagree because they do not refer to the same property with the term. As a consequence of the last difficulty, analytic reductionism is unable to account for normative divergence.

The realist explanation of persistent moral disagreements does not have these difficulties. First, it does not imply that some people involved in debates concerning rightness are linguistically incompetent in the sense of not knowing the meaning of ‘right’. According to the realist explanation, it can be assumed that people involved in such debates are linguistically

competent with respect to the term in so far as they know its meaning as it is stated in the realist formula. Second, it does not imply that people do not disagree because they do not refer the same property. According to the realist explanation, they refer to the same property with 'right', viz. to rightness as the property is characterised in the realist formula, and they may thus disagree in the relevant respect. The last point is especially important since it indicates that realism, unlike analytic reductionism, is able to account for normative divergence.

Moreover, the realist explanation is not vulnerable to Brink's argument against analytic reductionism. As Brink understands that view, it says that the meaning of a moral term consists in the set of non-moral properties that competent speakers conventionally associate with the term. Brink argues that it is possible to disagree with people who embrace the prevailing moral conception and consequently to reject the conception expressed in the assumed meaning of such a term. However, on analytic reductionism as Brinks understands it, this is not possible, since the meaning of a moral term consists in a set of non-moral properties of the mentioned kind. However, according to the realist explanation, this is possible, since the meaning of a moral term is not constituted by such a set of non-moral properties. Furthermore, according to the account offered by realism, people may disagree even if they embrace widely different moral conceptions.

Consider next the first kind of reductionist explanation when employed by *synthetic reductionism*. As we saw in section 7, synthetic reductionism has mainly three difficulties in relation to this explanation. First, it opposes our notion of what people who are involved in disagreements about normative conceptions refer to with 'right'. Second, it construes the content of certain sentences concerning normative conceptions in a way that opposes our notion of what these sentences say. Finally, contrary to what synthetic reductionism implies, people who embrace widely different normative conceptions may disagree even if their use of 'right' is not causally regulated by the same non-moral property. Particularly as a consequence of the last difficulty, synthetic reductionism fails to account for normative divergence.

I think it can be argued that the realist explanation does not have these difficulties.

Let us start with the first problem for synthetic reductionism mentioned above. The source of this difficulty is the view that people refer with a moral term such as ‘right’ to a non-moral property of the kind specified by a normative theory. Realists oppose this view because they believe that ‘right’ does not refer to such a non-moral property; indeed, in their view, the term does not refer to a particular non-moral property at all. To see the relevance of this difference between synthetic reductionism and realism to the problem under discussion, we may return to the example mentioned in section 7. In this example, ‘right’ refers to maximising happiness according to synthetic reductionism, and a utilitarian and a deontologist debate their normative conceptions. Synthetic reductionism implies then implausibly that the deontologist refers to maximising happiness with ‘right’. According to the realist explanation, this is not the case. Rather, the deontologist refers to rightness as the property is characterised in the realist formula. Roughly put, in asserting that an action is right, she claims that it has a set of non-moral properties such that it makes actions right. Moreover, she holds that the normative conception she embraces specifies a set of non-moral properties that makes actions right. Hence, the realist explanation does not have the untenable consequence of synthetic reductionism.

Let us next consider the problem that synthetic reductionism construes the content of certain sentences relating to normative conceptions in a way that opposes our notion of what these sentences say. This difficulty has the same source as the first one. Realism is not susceptible to this difficulty because, again, it does not take ‘right’ to refer to a non-moral property that is specified by a normative theory. To see the relevance of this difference between synthetic reductionism and realism, we may return to one of the examples mentioned in section 7. In this example, it was once more assumed that ‘right’ according to synthetic reductionism refers to maximising happiness. It was further assumed that a utilitarian utters the sentence ‘Actions that maximise happiness are right’. According to synthetic reductionism, this sentence says then, implausibly, ‘Actions that maximise happiness are such that they maximise happiness’. However, according to realism ‘right’ in the first sentence does not refer to maximising happiness or any other non-moral property that is specified by a normative theory. Rather, it refers to rightness as this property is characterised in the realist

formula. As a consequence, realists do not construe the sentence in the same way as synthetic reductionism does. Similar comments hold for the other types of sentences mentioned in section 7.

Finally, we saw that there is reason to believe that people may disagree about their normative conceptions even if these conceptions are so different that it is implausible to assume that their use of ‘right’ is causally regulated by the same non-moral property. On synthetic reductionism, this is not possible because people whose use of ‘right’ is not casually regulated by the same non-moral property do not refer to the same property and hence do not disagree in the relevant respect. On realism, however, people whose use of ‘right’ is not causally regulated by the same non-moral property may refer to the same property and can accordingly disagree. The reason is that on this view, ‘right’ does not refer to a non-moral property that causally regulates our use of the term; indeed, it does not refer to a particular non-moral property at all. Rather, it refers to rightness as this property is characterised in the realist formula.<sup>45</sup> That realism avoids this

---

<sup>45</sup> Put explicitly, the argument is the following. Given that two persons differ extensively in their normative conceptions and consequently differ to a significant extent concerning what they apply ‘right’ to, it seems implausible to assume that their use of the term is causally regulated by the same non-moral property. In particular, given that they differ extensively in the mentioned respect, it is implausible to assume that a certain non-moral property—such as maximising happiness in the example above—is causally responsible for certain reference-fixing characteristics that both of them utilise to apply the term. According to synthetic reductionism, it follows that they do not refer to the same property, rightness and so do not disagree. However, we do not seem prepared to accept that conclusion.

It might be wondered whether realism is not vulnerable to this objection as well. The reason to suspect this is that I proposed in chapter 4 that realists may adopt the view that we are able to apply a moral term such as ‘right’ correctly by help of reference-fixing characteristics that we associate with the term. However, there is an important difference between synthetic reductionism and realism. According to synthetic reductionism, a moral term such as ‘right’ refers to the non-moral property that is causally responsible for the reference-fixing characteristics and so causally regulates people’s use of the term. This means that if two persons do not have their use of ‘right’ causally regulated by the same non-moral property, they do not refer to the same property. However, on realism the reference-fixing characteristics are not connected to reference in this way. According to this view, it is not the non-moral property that ‘right’ *refers* to that is causally responsible for the reference-fixing characteristics; indeed, there is no such non-moral property on this view. Rather, according to realism, it is the non-moral properties that *make* actions right that are causally responsible for these characteristics. This means that people whose use of ‘right’ is not causally regulated by the same non-moral property may nevertheless refer to the same property. On realism, people refer to rightness in so far as they refer to this property as it is characterised in the realist formula. However, it should be admitted that if a person does not apply ‘right’ to actions that have the reference-fixing characteristics that we associate with the term to a certain extent, we would presumably doubt that she refers to rightness with the term. We might also be inclined to say that if she does not apply ‘right’ to actions that have certain of the reference-fixing



difficulty is especially important since it indicates that this view, in contrast to synthetic reductionism, is able to account for normative divergence.

Let us now turn to *the second kind of explanation* that reductionism employs.

According to this explanation, people disagree morally because of an error regarding non-moral facts. In section 8, I argued that the main difficulty with this explanation is that even if people involved in a moral disagreement take the moral term in question to have the same meaning or reference, they need not differ as regards any relevant non-moral beliefs. This is particularly likely to occur in relation to disagreements about normative conceptions and hence in relation to persistent moral disagreements.

The realist explanation of persistent moral disagreements is not vulnerable to this difficulty. According to this explanation, what explains persistent disagreements pertaining to rightness is an error as regards what makes actions right. More precisely, such disagreements are explained in terms of an error as regards an implication of the kind **(B)** in the realist formula: a synthetically necessary implication from a certain set of non-moral properties to rightness. This means that the antagonists in a persistent disagreement have conflicting beliefs about such an implication. What is important to note is that since such beliefs make reference to an irreducible moral property, they are not non-moral beliefs, but rather a kind of *moral* beliefs that concern a certain type of moral principle. Consequently, in contrast to what the second reductionist explanation implies, people who are involved in a disagreement concerning rightness may take the meaning or reference of 'right' to be the same and yet not differ as regards any relevant *non-moral* beliefs. Put differently, they may agree about the meaning or reference of 'right' and about all relevant non-moral beliefs; the only significant aspect in which they disagree is their beliefs about the mentioned kind of implication from a set of non-moral properties to rightness.

As we observed in section 8, one difficulty with the reductionist explanation of the second kind concerns the arguments presented in debates

---

characteristics, she is not linguistically competent with respect to the term. But it might still be the case that people who we believe refer to rightness differ in the way they apply the term—and so differ as to whether they apply it to actions that have the reference-fixing characteristics—to such an extent that it is implausible to assume that their use of the term is causally regulated by the same non-moral property.

about normative conceptions. Rather than being intended to point out that some non-moral facts are or are not the case, such arguments are typically intended to call attention to which non-moral properties contribute to, or fail to contribute to, actions being right. On the present proposal, a normative conception about rightness states which non-moral properties make actions right, in the sense of ‘make’ specified above. Accordingly, when arguments against normative conceptions appeal to the contribution of non-moral properties to the rightness of actions, realists may understand such arguments in the indicated sense of ‘make’. Recall the example mentioned earlier. On the account proposed by realism, the arguments put forward against utilitarianism are basically intended to call attention to either of two types of cases. Either that a set of non-moral properties makes actions right, but that utilitarianism does not accommodate this fact. Or that a certain set of non-moral properties does not, at least not always, make actions right, but that utilitarianism implies so; that is, it might be argued that what makes actions right is not that they maximise happiness, at least not always.

It should further be noted that the realist explanation of persistent disagreements avoids *relativism*. Unlike the version of relativism considered in section 9, realism does not imply that people involved in debates about normative conceptions use ‘right’ in such a way that its meaning or reference is constituted by different non-moral properties. Rather, on realism they use the term in such a way that its meaning and reference consists in rightness as this property is characterised in the realist formula. This means that realism does not have the untenable consequences of relativism.

It is important that the realist explanation of persistent moral disagreements appeals to a type of implication that is synthetically, not analytically, necessary. If the implication were understood in the latter way, it would be analytically necessary that if an action has a certain set of non-moral properties, it is right. In case such an implication holds, there is reason to believe that persons who are linguistically competent in the relevant respect and who have beliefs regarding the mentioned set of non-moral properties would also have beliefs regarding rightness. The reason is that if an implication is analytically necessary, linguistic competence involves knowledge of it. In that case it might not be tenable to claim that people

who are involved in persistent disagreements may share all relevant non-moral beliefs and yet disagree regarding rightness.

Finally, return to *widespread moral disagreements*. It might be argued that some of these disagreements are explainable in the same way as persistent moral disagreements. Suppose, for example, that people belonging to different cultures disagree to a great extent as to whether a certain kind of actions is right. According to realism, the reason might be that people in a certain culture embrace an erroneous normative conception. That is, they might be mistaken as regards what makes actions right.

## 12. *Explaining the Persistence of Persistent Moral Disagreement*

In the last section, I argued that realists can explain persistent moral disagreements in terms of an error concerning what makes objects have a certain moral property or, more precisely, in terms of an error concerning a kind of synthetically necessary implication from a certain set of non-moral properties to the moral property. However, it might be objected that this does not explain why these disagreements are persistent. It is correct that no explanation of the persistence of certain disagreements has been provided so far. However, the realist conception has some features that individually or collectively recommend such an explanation.

Firstly, it is important to bear in mind that persistent moral disagreements are explained in terms of an error regarding a kind of implication that is synthetically, not analytically, necessary. Had the kind of implication at issue been analytically necessary, people would most likely agree about it to a greater extent since they would have knowledge of it, in so far as they are linguistically competent in the relevant respect. By contrast, to get knowledge of a synthetically necessary implication may require quite thorough deliberation that involves various non-linguistic considerations. Hence, knowledge about such an implication may not be easy to come by and people could debate it extensively without reaching agreement.

Secondly, it might be argued that the question as to whether an implication of the mentioned kind holds raises intricate philosophical issues. As was argued above, disagreements regarding such implications are expressions of conflicting normative conceptions. Consider an example of such a conflict, the one between utilitarians and deontologists as regards

rightness. Which of these views, if any, that is correct might depend on a number of philosophical problems, e.g. the relation between moral rightness and other evaluative notions, the nature of personhood and the moral relevance of phenomena like friendship, love and integrity. Disagreements about other normative conceptions might be connected to other but nonetheless significant philosophical issues. If there is such a relation between normative conceptions and philosophical problems, it should not come as a surprise that disagreements concerning normative conceptions are persistent. However, this suggestion brings up epistemological problems that I will not discuss here.

Thirdly, it should be remembered that the realist formula means that there may be more than one set of non-moral properties that makes objects have a certain moral property ('multiple realisability'). There may in other words be more than one synthetically necessary implication of the mentioned kind. This fact can have the effect that discussions about moral matters become more complicated than they otherwise would have been, thus helping to explain why certain moral disagreements are persistent.

Finally, a set of non-moral properties that according to the realist formula makes something have a moral property might be immensely complicated, involving a long range of non-moral properties that are related to each other in various ways. A set of non-moral properties with such a complex character might presumably provoke disagreements that cannot easily be settled.

### **13. *Concluding Remarks***

In this chapter, I have considered an argument against the existence of moral properties that takes its point of departure in moral disagreement: the argument from moral disagreement. We have seen that if feature F in premise (2) of this argument is assumed to consist in moral disagreements being persistent, a significant version of this argument is obtained. Advocates of the existence of moral properties should then try to reject premise (3) of the argument thus understood. In order to do so, they should provide a plausible explanation of persistent disagreements in terms of an error in relation to a moral property. I have argued that realism, in contrast to reductionism, is able to provide such an explanation. An important reason is that realism, unlike reductionism, is able to account for the phenomenon of

normative divergence. As realism is able to provide a plausible explanation of persistent disagreements, it is able to refute premise (3) in the sense that opponents to the existence of moral properties—non-cognitivists and error-theorists—are not justified to claim that the best explanations of such disagreements do *not* imply that there is an error in relation to a moral property. Hence, they are not in the position to draw the conclusion that there are no moral properties. However, it can also be argued that realism is able to refute premise (3) in a more thorough way by maintaining that the best explanations of persistent disagreements *do* imply that there is an error in relation to a moral property. In section 4, I argued, against non-cognitivism and error-theory, that there are *prima facie* reasons to believe that the best explanations of moral disagreements are made on the assumption that there are moral properties. This means that the best explanations of such disagreements imply that there is an error in relation to a moral property. This suggests in turn that either reductionism or realism provides the best explanations of such disagreements. Now, since realism, in contrast to reductionism, is able to provide the required explanation, it is reasonable to claim that it provides the best explanations of the relevant kind of moral disagreements. Hence, there is reason to claim that it succeeds in rejecting premise (3) in the more thorough way just mentioned. Consequently, realism seems able to refute the argument from moral disagreement and hence defend the existence of moral properties against this argument. However, it should be recalled that there may be other significant arguments which take their point of departure in moral disagreement that I have not dealt with here.

*Appendix: Persistent Moral Disagreements and Unknowable Moral Facts*

There is another argument against the existence of moral properties that could also be interpreted as a version of the argument from moral disagreement. Like the version of the argument we have considered above, it can be understood to take the feature F in premise (2) to consist in moral disagreements being persistent. However, according to this version of the argument, the persistence of certain moral disagreements is taken as an indication of them being irresolvable ‘even in principle’. It is then claimed that if there were any moral properties, such disagreements would concern moral facts that are impossible to obtain knowledge of. On the assumption

that persistent disagreements are like this, it is argued, with premise (3), that the best explanations of such disagreements do not imply that there is an error in relation to a moral property.<sup>46</sup> It is then concluded that there are no moral properties.

This argument requires without doubt more attention than I am able to give it here. Especially, it requires a discussion of moral epistemology. However, to my mind it is less plausible than the versions of the argument from moral disagreement that have occupied us thus far. As far as I know, no one has made plausible the claim that persistent disagreements are irresolvable ‘even in principle’. And in case there are such disagreements, it might be questioned whether they concern moral facts that are unknowable.<sup>47</sup> Furthermore, it might perhaps also be questioned whether the existence of unknowable moral facts means that the relevant type of disagreements are not best explained in term of an error in relation to a moral property.

It might also be asked why the view that there are unknowable moral facts should be considered so untenable. One answer might be that there is no explanation of why such facts are unknowable.<sup>48</sup> However, on the realist account offered above, there might be such an explanation. I mentioned earlier that persistent disagreements concerning conflicting normative conceptions might raise complex philosophical problems. That such problems might concern facts that are unknowable is, I think, nothing we would regard as terribly upsetting, although disappointing. *If* there are unknowable moral facts, this might be a consequence of their being unknowable philosophical facts. However, I do not claim that there are any moral or philosophical facts that are unknowable.

There are also other considerations that may make us reluctant to the view that there are unknowable moral facts. One such consideration is that this view would suggest that many, perhaps all, moral facts are unknowable. Another consideration is that it would be unreasonable to hold that some moral facts that we take to be evident are unknowable.<sup>49</sup> However, the

---

<sup>46</sup> See e.g. Bennigson (1996), pp. 411–437. Wright has put forward an argument which reminds of this reasoning; see Wright (1992), pp. 140–157. For a forceful criticism against this kind of argument, especially as it is formulated in Wright, see Tersman (2002), chap. 3.

<sup>47</sup> Cf. Tersman (2002), pp. 68–76.

<sup>48</sup> See e.g. Wong (1984), p. 152.

<sup>49</sup> See e.g. Bennigson (1996), p. 413.

realist account avoids these consequences. What would be unknowable on the proposal above is which normative conception is correct; that is, whether certain non-moral properties make objects have a certain moral property. That this is unknowable is however compatible with most other kinds of moral facts being knowable. In chapter 4, I suggested that we might be able to apply ‘right’ correctly, and so pick out the actions that are right, by means of reference-fixing characteristics that we associate with the term. In order to do so, we need not know—at least not exactly—what non-moral properties make actions right. So, even if we cannot come to know exactly what makes actions right, it does not mean that we cannot come to know many moral facts or that we cannot know the moral facts that we take to be evident. This account seems to gain support from moral phenomenology. We do not think we have to know what non-moral properties make actions right—at least not exactly—to be justified to claim that a certain action is right. What we with certainty take ourselves to know is that there *are* some non-moral properties that make them right.

## Chapter 6

### Explaining Moral Reason

#### 1. Introduction

In the meta-ethical literature, it is frequently pointed out that a distinguishing feature of moral judgements is that they involve reasons to perform actions. One of the strongest arguments against a meta-ethical view is accordingly thought to be that it fails to account for this feature of moral judgements. However, the claim that moral judgements involve reasons is ambiguous. As is often pointed out, one should distinguish between two kinds of reasons for performing actions.<sup>1</sup> There is a *normative reason* to perform a certain action if there is a norm or standard for assessing actions that generates the reason. If a person performs an action there is a normative reason to do, her action is justified from the perspective of the standard in question. There is a *motivating reason* to perform a certain action if a person who has such a reason is motivated to perform the action. If a person performs an action that she has a motivating reason to do, the reason can be appealed to in an explanation of her action. In view of this distinction, there are basically two ways in which moral judgements can be claimed to involve reasons: by involving normative reasons and by involving motivating reasons. In the present chapter, I will discuss the relation moral judgements have to the first kind of reasons, and in the next chapter I will discuss the relation they have to the second kind of reasons. That moral judgements involve normative reasons is often put by saying that they are *normative* or that they have *normativity*. Henceforth when I refer to ‘reasons’ in this chapter, I have in mind normative reasons. In both chapters, I will be concerned with moral judgements pertaining to rightness.

In one fundamental sense it is uncontroversial that judgements to the effect that actions are morally right involve normative reasons, namely in so far as they imply *moral* reasons. What is controversial is what moral reasons consist in, particularly whether they consist in reasons of rationality, as rationalism claims. This is presumably the issue most often discussed in the meta-ethical literature on reasons. However, it is not the topic of the

---

<sup>1</sup> See e.g. Cullity and Gaut (1997), pp. 1–3, and Smith (1994), pp. 95–96.



present chapter. Rather, I will be concerned with a more basic issue in that I will consider the relation between the moral facts stated by moral judgements and the moral reasons implied by such judgements. I will argue that reductionism and realism construe this relation differently and that this has consequences for the plausibility of these views. Reductionism, I will argue, construes this relation in such a way that it is unable to answer fundamental normative questions and to provide an accurate account of the normativity of moral judgements. However, I will argue that realism construes this relation in such a way that it is able to answer these questions and to provide a reasonable account of the normativity of such judgements. In contrast to the previous chapter and the next one, I will in this chapter not discuss an argument that has been taken to support non-cognitivism or error-theory, but an argument with relevance for the plausibility of reductionism and realism, respectively.

In the next section, I introduce what I call ‘the moral reason principle’ which says, roughly put, that a moral judgement to the effect that an action is morally right entails that there is a moral reason to perform that action. The subsequent two sections concern how reductionism and realism, respectively, read the moral reason principle. In section 5, I argue that because of the way reductionism reads this principle, it fails to provide a satisfactory answer to a version of the question ‘Why should I do what is right?’ However, realism reads the principle in such a way that it is able to provide a reasonable answer to this question. This difference indicates, I maintain, that realism, unlike reductionism, is able to provide a plausible account of the normativity of the pertinent moral judgements. In section 6, I respond to three objections that might be directed against the argument presented in the previous section. A number of philosophers have argued on similar lines regarding analogous normative questions, and in section 7 I argue that their arguments can be understood in accordance with my earlier reasoning. In section 8, I discuss some other normative questions. In section 9, I argue briefly that there are reasons to believe that the realist reading of the moral reason principle is compatible with rationalism. However, this issue deserves a more thorough discussion.

## 2. *The Moral Reason Principle*

Above we saw that there is a reason—i.e. a normative reason—to perform an action if there is a norm or standard for assessing people’s actions that generates it. Among others Philippa Foot has pointed out that there are numerous standards for assessing people’s behaviour that generate such reasons, e.g. etiquette, law, morality, prudence and rationality.<sup>2</sup> Consequently, there are e.g. reasons of etiquette, juridical reasons, moral reasons, prudential reasons and reasons of rationality to perform or not to perform actions. Here I would like to draw attention to the relation between, on the one hand, judgements to the effect that actions are right and, on the other hand, reasons generated by such standards. When it is judged that an action is right, it is normally implied that the action is right according to a certain standard for assessing people’s behaviour. Now, the following seems to be the case: on any such standard, if an action is right according to the standard, then there is a reason, according to that standard, to perform the action. A corresponding principle holds for wrongness. We may take an example from etiquette. Suppose someone remarks ‘It’s wrong to lick on one’s knife while eating’ and ‘wrong’ is to be understood as ‘wrong according to etiquette’. If this judgement is correct, it follows that according to etiquette there is a reason—an etiquette reason—not to lick on one’s knife while eating. We can ask an expert on etiquette: ‘Why is it wrong to lick on one’s knife while eating?’ She is likely to reply with what she takes to be an etiquette reason. She might reply, for example, ‘It’s impolite’. However, if she answers that there is *no* such reason, we would doubt that licking on one’s knife while eating actually is wrong according to etiquette, granted that we trust her knowledge of these matters.

Morality is a standard for assessing people’s behaviour that exhibits the observed connection between judgements to the effect that actions are right and reasons. Sometimes we use ‘right’ to mean ‘right according to morality’ or, simply put, ‘morally right’. Like other standards for assessing people’s behaviour, morality generates reasons for action, reasons according to morality or, shortly put: moral reasons. Thus, morality provides us with moral reasons to perform certain actions. Now, it seems uncontroversial to claim that if an action is morally right, it follows that there is a moral reason

---

<sup>2</sup> Foot (1978 (1972)), pp. 159–161. Cf. Brink (1992), pp. 8–9; Brink (1997), pp. 20–21; Smith (1994), pp. 95–96, and Smith (2001a), p. 119.

to perform that action. It seems also to be the case that to know that ‘right’ is to be applied in conformity with this notion is a condition for being linguistically competent with respect to the term. Suppose a person claims that a certain action is morally right, but that she also believes that there is no moral reason to perform it. I think we would regard this as an indication of her not being linguistically competent in respect of the term. It seems consequently reasonable to claim that the following principle holds:<sup>3</sup>

*The moral reason principle:* It is analytically necessary that if it is morally right to perform an action, then there is a moral reason to perform that action.<sup>4</sup>

An analogous principle holds for ‘morally wrong’. In what follows I will omit ‘morally’ in ‘morally right’ as I throughout will be concerned with moral rightness.

In the present context, a reason to perform an action is understood as a fact to the effect that the action has a certain property, F. So understood, a reason consists in the fact *that* the action has F.<sup>5</sup> Such a property can be said to constitute the reason. When a person performs an action that she has this reason to perform, she consequently performs an action that has this property, F. According to the moral reason principle, if an action is right, it follows that there is a moral reason to perform it. The moral reason will consequently consist in the fact that this action has a certain property.

As already indicated, it seems incontrovertible that a moral judgement to the effect that an action is right entails the existence of a moral reason in the way stated by the moral reason principle. In the primary sense given by this principle, it is thus uncontroversial that such moral judgements are normative. Something that *is* controversial is what the moral reason referred to in this principle consists in and, hence, what the normativity of such moral judgements amounts to. According to a family of views, the reasons morality provides should be understood in terms of reasons generated by

---

<sup>3</sup> Cf. Brink (1992), pp. 1, 8–9; Brink (1997), pp. 20–21; Copp (1995), p. 190; Lillehammer (1999), p. 121; Railton (1993a), pp. 295–296; Shafer-Landau (2003), p. 166, and Smith (2001a), p. 119.

<sup>4</sup> It might sometimes be appropriate to insert the qualification that an action is right *for* a person to perform. However, as this complication is not relevant to my arguments in the present chapter, I will omit the qualification.

<sup>5</sup> See e.g. Raz (1975), pp. 16–20. In what follows, I will presume that the reason in question is a complete reason to perform a certain action, e.g. a complete moral reason to perform a certain action. Often when people state what they take to be a reason to perform an action, they only mention a part of such a reason. For an account of the notion of complete reasons, see Raz (1975), pp. 22–25.

some of the other normative standards for assessing people's behaviour, e.g. prudence or rationality. Thus understood, moral reasons comprise a subclass of the reasons belonging to this kind of normative standard and moral reasons consequently consist in e.g. prudential reasons or reasons of rationality.<sup>6</sup> (In what follows, I will refer to the latter kind of reasons as 'rational reasons'.) According to rationalism, morality provides rational reasons, and moral reasons consist consequently in reasons of that kind. This view yields a principle which says that, necessarily, if an action is right, there is a rational reason to perform it. Judgements to the effect that actions are right would then be normative in the strong sense that they imply that there are rational reasons to do what is right. Needless to say, there are various views as to what rational reasons consist in, and what moral reasons consist in if they do not consist in such reasons.<sup>7</sup>

As mentioned earlier, I do not intend to argue for or against any particular view of the type just mentioned as regards what moral reasons consist in. Moreover, I think that, as far as my arguments are concerned, reductionism and realism are possible to combine with various views about such reasons. The subsequent discussion will instead concern the more formal issue of how the moral reason principle should be interpreted internally. More precisely, I have in mind the relation between the facts referred to in the antecedent and the consequent of this principle: on the one hand, the fact that an action is right, and, on the other hand, the moral reason. According to the reading to which reductionists are committed, these facts are identical: the moral reason consists in the fact that the action is right. According to another reading, open to realists, these facts are not identical: the moral reason does not consist in the fact that the action is right, but in the fact that the action has a set of non-moral properties which makes actions right. This difference, I will argue, is relevant to the plausibility of reductionism and realism, respectively.

---

<sup>6</sup> Cf. Brink (1992), pp. 8–9; Brink (1997), pp. 20–21, and Smith (1994), pp. 95–96. See also Darwall (1990), pp. 257–259.

<sup>7</sup> For helpful overviews of different theories about moral reasons, see Brink (1992), pp. 1–26; Cullity and Gaut (1997), pp. 1–27, and Parfit (1997), pp. 100–105.

### 3. *Reductionism on the Moral Reason Principle*

According to reductionism, rightness is identical with a non-moral property G. This view of the relation between rightness and non-moral properties has implications for how reductionism reads the moral reason principle.

In the last section, we saw the moral reason principle says that it is analytically necessary that if it is right to perform a certain action, there is a moral reason to perform that action. Now, I think it is reasonable to argue that because of their conception of the relation between rightness and non-moral properties, reductionists are committed to the view that the moral reason is identical with the fact that the action is right. To illustrate, consider a version of reductionism according to which the non-moral property G consists in maximising happiness, i.e. a version of reductionism that identifies rightness with maximising happiness. A reductionist advocating this view would then have to claim that the moral reason at issue consists in the fact that the action has G, i.e. the fact that the action maximises happiness. Since maximising happiness on this version of reductionism is identical with rightness, the reductionist identifies the moral reason with the fact that the action is right.

That this reading of the moral reason principle correctly represents the relation between moral properties and moral reasons is something that often seems to be taken for granted in the meta-ethical literature. For example, it is sometimes said or implied without argument that the fact that an action is right *is* a reason to perform the action. Often it is claimed, more vaguely, that the fact that an action is right provides or gives a reason to perform the action. Accordingly, ‘right’ is sometimes used interchangeably with terms such as ‘moral reason’, ‘reason’ and other terms that refer to such a reason, e.g. ‘should’. Some authors are also explicit about their acceptance of this conception of the relation between rightness or some other moral property and moral reason.<sup>8</sup> It is significant that they generally do not seem to think that this conception needs to be argued for.<sup>9</sup>

---

<sup>8</sup> See e.g. Smith (1994), pp. 182–184. For related views, see e.g. Jackson (1998), pp. 141–142; Johnston (1989), p. 154; Mackie (1977), pp. 73–80, and Scanlon (1998), pp. 11–12.

<sup>9</sup> It might be argued that this does not hold according to versions of reductionism that are subjectivist. For example, consider a subjectivist version of reductionism according to which the rightness of actions consists in their being approved of by a person. It might be claimed that the moral reason on this view does not consist in the fact that an action is approved of by a person and hence not in the fact that an action is right. Rather, the moral reason consists in the fact that the action has certain features, viz. the features that

Although it seems reasonable to assume that the mentioned conception is the received view, it is worthwhile to make explicit why reductionism is committed to it.

According to the moral reason principle, it is analytically necessary that if an action is right, there is a moral reason to perform that action. In the last section, we saw that the moral reason consists in the action having a certain property. Suppose a person performs the action that there is this reason to perform. She then performs an action that is right. It should next be observed that on the reductionist view that rightness is identical with a non-moral property G, this moral reason has to consist in the fact that the action has G. As we have seen, if a person performs an action that there is a moral reason to perform, and the reason is entailed according to the moral reason principle, she has performed an action that is right. According to the necessary biconditional to which reductionism is committed, it is only if the moral reason consists in the fact that the action has G that this holds. It is only if the moral reason consists in the fact that the action has G that it follows that if a person performs an action that there is this reason to perform, she performs an action that is right. Now, reductionism identifies rightness with G. Hence, it is committed to the view that the moral reason consists in the fact that the action is right.<sup>10</sup>

---

elicit the person's approval. One motive for this claim is that on the person's own view, these features presumably constitute her reason for why she approves of the action, whereas the attitude itself does not form part of her reason. However, I think that also on subjectivist versions of reductionism, the moral reason should be understood according to the first alternative. What the person in question takes to be her reason to perform an action should be distinguished from what constitutes the moral reason to perform an action. On the person's own view, her attitude is presumably not part of her reason. However, the approval is presumably part of the moral reason. One ground for this view is that if the person had not had the attitude in question towards the action, there would not be a moral reason to perform that action according to the view under consideration. Thus, if we inquire what the moral reason is to perform an action according to this version of reductionism, we would have to refer to a person's approval to correctly describe the reason. The arguments below can be taken to confirm this view.

<sup>10</sup> To illustrate, return to the example of a version of reductionism mentioned above according to which G consists in maximising happiness. On this view, the moral reason mentioned in the moral reason principle has to consist in the fact that the action maximises happiness. This is so since, according to the necessary biconditional to which this version of reductionism is committed, it is only if the moral reason is understood in this way that it follows that, if a person performs the action that there is this reason to perform, she performs an action that is right. Since rightness is identified with maximising happiness on this view, the moral reason has accordingly to consist in the fact that the action is right.

Another way to reach this conclusion is by means of a *reductio*. Assume that there is a moral reason that is entailed in accordance with the moral reason principle which does *not* consist in the fact that the action has G. Suppose a person performs an action that there is this moral reason to perform and hence performs an action that is right. On the mentioned assumption it would then be possible to perform an action that is right without performing an action that has G. But then rightness would not be identical with G and reductionism would be false. Hence, in order to reduce rightness to G, reductionists are committed to the view that the moral reason that an action's rightness entails consists in the fact that the action is right.<sup>11</sup>

#### 4. *Realism on the Moral Reason Principle*

According to realism, the relation between rightness and non-moral properties is described in the realist formula, and this principle states that rightness is not identical with any non-moral property. This view of the relation between rightness and non-moral properties has consequences for how realism reads the moral reason principle.

According to the moral reason principle, it is analytically necessary that if an action is right, there is a moral reason to perform it. As we saw in the last section, reductionists are committed to the view that the moral reason is identical with the fact that the action is right. This is so because reductionists claim that rightness is identical with a non-moral property. Since realists deny this view, they are free to oppose the mentioned reading of the moral reason principle. Realists may instead adopt the following reading. On this conception, the moral reason of the relevant kind consists in the fact that the action has a set of non-moral properties G, where this set is such that it *makes* the action right. That is, the moral reason entailed by the fact that an action is right consists in the fact that the action has a set of non-moral properties of the kind mentioned in the realist formula.<sup>12</sup>

---

<sup>11</sup> Return to the example of a version of reductionism above. Assume that the moral reason mentioned in the moral reason principle does *not* consist in the fact that an action maximises happiness. Suppose a person performs an action that there is this moral reason to perform and hence performs an action that is right. On the mentioned assumption, it would then be possible to perform an action that is right without performing an action that maximises happiness. But then rightness would not be identical with maximising happiness and this version of reductionism would be mistaken.

<sup>12</sup> Since there might be a number of sets of non-moral properties that make actions right according to realism, there might also be a number of moral reasons on this view.

According to the realist formula, rightness is not identical with such a set of non-moral properties. Consequently, the moral reason is not identical with the fact that the action is right.<sup>13</sup>

We may illustrate the difference between reductionism and realism by means of an example. Recall the version of reductionism mentioned above which says that rightness is identified with maximising happiness. Consider a realist counterpart to this view according to which a set of non-moral properties of the kind mentioned in the realist formula consists in maximising happiness. According to this version of realism, although maximising happiness is not identical with rightness, maximising happiness makes actions right. Consequently, a moral reason entailed in accordance with the moral reason principle consists in the fact that the action maximises happiness. As maximising happiness is not identified with rightness, the moral reason entailed by the fact that an action is right does not consist in the fact that the action is right.

##### 5. ‘*Why Should I Do What Is Right?*’

We often ask what reasons we have to perform certain actions, a type of normative questions that we frequently put by asking *why we should* do so and so. Such normative questions may also be asked in relation to moral issues. Often these questions ask about a reason to perform one particular action; we can for example ask ‘Why should I give him the money back?’ However, sometimes normative questions concerning moral issues ask about a reason to perform a certain class of actions which has one particular property. A fundamental question of this kind is:

‘Why should I do what is right?’

To make explicit that this question asks about a reason, we may reformulate it in the following way:

‘What reason is there for me to do what is right?’<sup>14</sup>

---

<sup>13</sup> According to the moral reason principle, it is analytically necessary that if an action is right, there is a moral reason to perform that action. This means, as we have seen, that if a person performs an action that there is this reason to carry out, she performs an action that is right. In the last section, we saw how reductionism accounts for this. Realism accounts for this in the following way. On this view, the moral reason consists in the fact that an action has a set of non-moral properties which is such that it makes actions right. If an action has such a set of non-moral properties, it follows according to the realist formula that it is right. Consequently, if a person performs an action that there is this moral reason to perform, she performs an action that is right. In contrast to what reductionism implies, the moral reason is not identical with the fact that the action is right.



In this section, I will argue that reflections on a version of this question suggest that the reading realism suggests of the moral reason principle is preferable to the reading suggested by reductionism. However, before we consider this version of the question, it is important to get clear over what questions of this type amount to. That is, we have to consider what the purport is of questions of the form ‘What reason is there to do what is F?’, where ‘F’ refers to a certain property. As far as I understand, there are basically two readings of this type of questions.

According to the *first interpretation*, a question of this type asks what reason there is to perform a certain class of actions, one of the properties of which is F. On this interpretation, it is not of any special importance that these actions have this particular property rather than some other one. To ask about a reason to do what is F is just one way to ask about a reason to perform this class of actions. This class of actions could be picked out by means of some other property the actions have in common. Instead of asking what reason there is to do what is F, it could then just as well be asked what reason there is to do what has that other property.

I think this interpretation is flawed for at least two reasons. To illustrate this, we might consider a non-moral example of the kind of questions under consideration: ‘What reason is there to do what is legal?’<sup>15</sup>

Firstly, the first interpretation construes questions that actually ask about different reasons to be asking about the same reason. Consider our example: ‘What reason is there to do what is legal?’ Suppose the class of actions that are legal also have another property in common. For example, assume that these actions also happen to be such that they would be approved of by a certain person called Bill. On the first interpretation, it is not of any special importance that these actions are legal rather than having some other property. This means, as we have seen, that to ask about a reason to do what is legal is just one way to ask about a reason to perform these actions and that one instead could ask about a reason to do what has some other property these actions have in common. According to this

---

<sup>14</sup> In what follows, I will omit the phrase ‘for me’, as I do not think it is relevant to the present discussion. It might be thought that the question ‘Why should I do what is right?’ can be understood in the following way: ‘Do I have a reason to do what is right?’ I will return to this reading in section 8.

<sup>15</sup> It is not significant here which kind of reason the question asks about. We have to presume, however, that it asks about the same kind of reason in the various contexts in which it is discussed below.

interpretation, instead of asking ‘What reason is there to do what is legal?’, it could then just as well be asked ‘What reason is there to do what Bill would approve of?’ That is, on the first interpretation, these questions ask about the same reason. However, this view seems mistaken, because these two questions ask about different reasons.

The following objection might be raised against this argument. In asking ‘What reason is there to do what is legal?’, a person need not know that the class of actions that are legal also happen to be such that they are approved of by Bill. It might then be suspected that *if* she knew this, she would take this question to ask about the same reason as ‘What reason is there to do what Bill would approve of?’ This objection seems mistaken, however, because even if we imagine a person who knows this, she would take these questions to ask about different reasons. Thus, suppose a person first asks ‘What reason is there to do what is legal?’ and then comes to know that legal actions also happen to be such that Bill would approve of them. She would still, after having gained this knowledge, take the original question to be asking about another reason than ‘What reason is there to do what Bill would approve of?’ This is so because in asking the original question, she explicitly focuses on actions being *legal* rather than on their having some *other* property. She asks what reason there is to perform actions that have *this* specific property, rather than some other one. Otherwise formulated, she asks what reason there is to perform actions in special consideration of their having a *particular* property: that of being legal. However, the question ‘What reason is there to do what Bill would approve of?’ does not capture the focus on actions being legal in her original question. On the contrary, in the other question, the focus is on quite another property. Consequently, it asks what reason there is to perform actions that have the property of being such that Bill would approve of them, rather than some other property. Put in another way, it asks about a reason to perform actions in special consideration of their being such that Bill would approve of them. Obviously, this is not the reason the person asks about.

Secondly, the first interpretation is mistaken because it takes questions to have the same answer that do not have the same answer. Suppose it is asked ‘What reason is there to do what is legal?’ A possible answer to this

question might be: ‘One avoids to be punished’.<sup>16</sup> Suppose, again, that the class of actions that are legal happen to be such that Bill would approve of them. On the first interpretation, instead of asking the mentioned question, it could instead be asked ‘What reason is there to do what Bill would approve of?’ Since the first interpretation implies that these questions ask about the same reason, it suggests that they have the same answer, e.g. the one just indicated. However, this seems incorrect. The indicated answer might provide a plausible answer to the former question, but it does not seem to provide a plausible answer to the latter one.

According to the *second interpretation*, a question of the type ‘What reason is there to do what is F?’ does not ask what reason there is to perform a class of actions, one of the properties of which happens to be F. Rather, as already has been indicated, it asks what reason there is to perform actions that have the particular property F, rather than some other property. That is, it asks about a reason to perform actions in special consideration of their having the particular property F. This means, at variance with the first interpretation, that even if the class of actions that is F also has some other property in common, the question cannot be understood to ask what reason there is to do what has that property, since this would be to ask about quite another reason. The second interpretation can be clarified in the following way. According to this interpretation, the question asks what reason there is to perform actions that are F *irrespective* of which *other* properties they have. The question could then be understood to ask what reason there would be to perform these actions even if they had not had any other property than F. However, we should make exception for those properties that are needed for something to be F and those that are consequences of something being F, since these properties might be relevant to the reason in question.

The correctness of the second interpretation is confirmed by the fact that it does not have the difficulties of the first interpretation.

Firstly, the second interpretation takes questions that ask about different reasons to actually be doing so. Consider our example: ‘What reason is there to do what is legal?’ Suppose again that the class of actions that are legal happen to have another property in common, that they are such that they would be approved of by Bill. According to the second

---

<sup>16</sup> This might be a reasonable answer if the question is understood to ask about a prudential reason.

interpretation, the mentioned question and the question ‘What reason is there to do what Bill would approve of?’ do not ask about the same reason. This is so since they ask about reasons to perform these actions in special consideration of their having quite different properties: that of doing what is legal and that of doing what Bill would approve of, respectively. More exactly, whereas the first question asks what reason there is to perform actions that are legal, irrespective of which other properties these actions have (with the exceptions mentioned above), the second question asks what reason there is to perform actions that are such that Bill would approve of them, irrespective of which other properties these actions have (with the exceptions mentioned above). As can easily be seen, these questions do not ask about the same reason.

Moreover, the second interpretation takes questions that have different answers to actually having different answers. Suppose once more that the class of legal actions also happen to be such that Bill would approve of them. As we have seen, according to the second interpretation, the question ‘What reason is there to do what is legal?’ asks about another reason than the question ‘What reason is there to do what Bill would approve of?’ Consequently, these questions do not have the same answer. For example, although it might be plausible to answer the first question in the way indicated above, this would not be a plausible answer to the second question.

Consequently, questions of the type ‘What reason is there to do what is F?’, where ‘F’ refers to a certain property, should be understood according to the second interpretation.

Next we should notice that the question ‘Why should I do what is right?’ or ‘What reason is there to do what is right?’ can be understood in different ways depending on what kind of reason the question is taken to ask about. As we saw in section 2, there are numerous standards for assessing people’s behaviour that generate reasons for action, e.g. etiquette, law, morality, prudence and rationality. Two questions that are discussed in moral philosophy are ‘What rational reason is there to do what is right?’ and ‘What prudential reason is there to do what is right?’ One might perhaps also ask ‘What juridical reason is there to do what is right?’ There might also be other kinds of reasons with regard to which one might ask questions of

this kind. One might perhaps even ask ‘What reason of etiquette is there to do what is right?’, although this question might be peculiar.

After these clarifications, I would like to turn to a version of the question ‘What reason is there to do what is right?’ that is relevant to whether we should prefer the reductionist or the realist reading of the moral reason principle. In asking this question, we may take ‘reason’ to refer to a *moral* reason. That is, we may ask:

‘What moral reason is there to do what is right?’

It should be recalled that in referring to ‘right’, I have in mind moral rightness.

It may first be observed that our impression of this question is that it makes sense. Above it was implied that it is sensible to ask what reason there is to do what is right and have in mind e.g. a prudential reason or a reason of rationality. We do not seem to respond to the question under consideration in any other way; it seems that we just as sensibly can ask what reason there is to do what is right and have in mind a *moral* reason. Since the question makes sense, it is reasonable to assume that it has an answer.

Another way to see that the question ‘What moral reason is there to do what is right?’ makes sense is the following. As was observed in section 2, according to a family of views, the reasons morality provides should be understood in terms of the reasons generated by some of the other normative standards for assessing people’s behaviour, e.g. prudence or rationality. For instance, according to rationalism, morality provides rational reasons, and moral reasons consist consequently in rational reasons. On this view, it holds necessarily that if an action is right, there is a rational reason to perform that action. Advocates of this view investigate what such a reason to do what is right consists in; that is, they investigate what a moral reason that is identified with a rational reason consists in. They ask consequently what rational reason there is to do what is right. As far as I see, this question makes perfectly good sense; indeed, it has to make sense if such an investigation is to be worthwhile. Now, since the relevant moral reasons and rational reasons are identical according to the mentioned view, this means that if this question makes sense to ask, so does the question under discussion here.

Next it should be observed that ‘What moral reason is there to do what is right?’ is a fundamental normative question and that it is important that it can be answered. One way to illustrate this is to compare it with other normative questions that ask about moral reasons. Above we observed that we may ask normative questions about moral issues. Like the question under discussion, they are often plausibly understood to be asking about a moral reason. There are various questions of this kind. Above I suggested that normative questions about moral issues may inquire about a moral reason to perform a certain action, such as ‘Why should I give him the money back?’ Understood as a question about a moral reason, it becomes ‘What moral reason is there for me to give him the money back?’ Furthermore, as we have just seen, we can ask normative questions about moral issues where we ask about a reason to perform a class of actions in special consideration of their having a particular property, such as ‘Why should I do what is legal?’ Sometimes these questions ask about a moral reason. Thus, if we understand the mentioned question to be asking about a moral reason, it becomes ‘What moral reason is there to do what is legal?’

In comparison with these questions, ‘What moral reason is there to do what is right?’ asks a more fundamental normative question. In the situation in which it is asked, a person recognises that morality classifies certain actions as right, i.e. morally right. The person then asks what moral reason there is to do what is right. It can thus be understood as one of the ultimate normative questions in relation to morality, since it asks what reason morality itself provides us with to perform actions in consideration of their having the very property of being right. This is a question that is asked in real life and in a number of different situations. Moreover, it is a question that we think is of vital importance that it can be answered, especially as it tends to force itself on us in certain significant situations in our lives.

One type of cases where this question might arise are those where we have started to distrust morality. A person who has started to do so might end up with adopting a certain kind of moral scepticism with the consequence that she will cease to try to do what she believes is right. As an illustration, suppose that a person—perhaps after reading Marx and Engels—has come to suspect that the classification of certain actions as right is put into practice merely as a means for the upper classes to prevent the lower classes from making revolution. However, she has not yet quite made up

her mind whether this actually is so. She might then ask what reason morality as a matter of fact provides to do what is right. That is, she might ask what moral reason there is to do what is right—what a moral reason to do what is right actually consists in. The answer she arrives at might be decisive for whether she becomes a moral sceptic of the indicated kind and accordingly for whether she will continue to try to act rightly.

A second type of cases are those where we ask this question while comparing alternative ways of leading our lives and the actions these alternatives would demand from us. This is something we might do when we go through existential crises or face fundamental choices in life. One alternative a person who finds herself in such a situation may contemplate is to live in accordance with morality, something that would require her to perform actions that are right. As a part of the process of comparing this alternative to other ways of living and the actions these alternatives would involve, she might ask what reason morality provides to do what is right—what moral reason there is to do what is right. The answer she reaches may have implications for which kind of life she ultimately decides to live.

A third type of cases are those where we ask this question as a part of our deliberation whether to perform certain actions. For example, suppose a person is convinced that there are certain actions that would be right for her to perform, but she also realises that these actions will require great sacrifices from her. As a result, she has great difficulties in deciding what to do. It might also be the case that her ordinary ways of moral deliberation do not yield an answer in the situation in which she finds herself. In an effort to decide what to do, she might start to reflect quite generally on the strength of the various considerations that might speak for and against doing what is right. As a part of this deliberation, she might then ask what reason morality provides to do what is right—what moral reason there is to do what is right. The answer she arrives at may have implications for what she ultimately decides to do.

Let us now consider how reductionism answers the question ‘What moral reason is there to do what is right?’ The reply reductionism gives is, in effect, ‘It’s right’. According to the way reductionism reads the moral reason principle, the moral reason at issue consists in the fact that an action has a non-moral property G with which rightness is identified. Consider the example of a version of reductionism mentioned above according to which

G consists in maximising happiness. When an advocate of this version of reductionism is asked ‘What moral reason is there to do what is right?’, her answer will be ‘The moral reason to do what is right is that it maximises happiness’ or, for short, ‘It maximises happiness’. Since rightness and maximising happiness are identical on this view, she answers in effect ‘The moral reason to do what is right is that it is right’, or, for short, ‘It’s right’. Similar results hold for other versions of reductionism.

We may now consider whether this answer to the question states a moral reason to do what is right. Consider a person who asks the question under discussion and gets an answer to the effect ‘It’s right’. Our reaction to such an answer is, I think, that it just repeats what is asked for and that it therefore cannot state the requested reason. A moral reason to do what is right cannot, properly understood, consist in the very fact that it is right.

It might be wondered what the underlying principle is that explains why we find the answer reductionism gives to the question under consideration unsatisfactory. I would like to offer, very briefly, one proposal as to what this principle is. In section 2, I said that in the present context a reason consists in a fact to the effect that the action has a certain property. Such a property, I suggested, constitutes the reason. It seems natural to think of the property that constitutes a reason as that which gives the reason its normative force, as the aspect of the reason that speaks in favour of performing the action in question. As noted above, we sometimes ask questions of the form ‘What reason is there to do what is F?’, where ‘F’ refers to a certain property. It is this kind of reason we ask about when we pose the question ‘What moral reason is there to do what is right?’ A principle that explains our response to the answer reductionism gives to this question might now be suggested. The principle is the following: a proposed reason to do what is F is such a reason only if the property that constitutes it is not identical with F. It is not difficult to come to think of a rationale of this principle. Unless a proposed reason fulfils this principle, it does not have any normative force. The explanation is that the property that constitutes the proposed reason—and hence is supposed to provide normative force to the reason—would not be distinct from what the proposed reason is meant to be a reason to do. Now, it is evident that what according to reductionism is a moral reason to do what is right does not satisfy this principle. On this view, the property that constitutes the moral



reason to do what is right is the very property of being right. Hence, what according to reductionism is a moral reason to do what is right does not have any normative force. According to the mentioned principle, the account reductionism offers therefore does not state a moral reason to do what is right. It seems reasonable to assume that this explains why we find the answer reductionism gives to the question under discussion unsatisfactory.

We may thus conclude that the answer reductionism gives to the question ‘What moral reason is there to do what is right?’—i.e. ‘It’s right’—does not supply the requested reason. That is, the judgement ‘It’s right’, as reductionism understands it, fails to state a moral reason to do what is right. It is, as we have seen, the way reductionism reads the moral reason principle that is responsible for this failure. There is therefore reason to claim that this reading of the moral reason principle is mistaken. As mentioned in section 2, the moral reason principle concerns the normativity of judgements to the effect that actions are right. Since the reading reductionism provides of this principle is erroneous in the indicated way, it also seems reasonable to assume that reductionism provides a flawed account of the normativity of such moral judgements.

Let us now consider how realism answers the question ‘What moral reason is there to do what is right?’ According to the realist reading of the moral reason principle, a moral reason of the relevant kind consists in the fact that an action has a set of non-moral properties *G* which is such that it makes actions right. Since such a set of non-moral properties is not identical with rightness, realism provides another answer than reductionism. We may illustrate this response with help of the example of a version of realism mentioned above according to which a set of non-moral properties that makes actions right consists in maximising happiness. On this version of realism, the moral reason at issue consists in the fact that an action maximises happiness. When someone asks an advocate of this version of realism ‘What moral reason is there to do what is right?’, she will reply ‘A moral reason to do what is right is that it maximises happiness’ or, for short, ‘It maximises happiness’. Since maximising happiness is not identical with rightness on this view, the realist is not committed to giving the answer that the reductionist gives. Similar results hold for other versions of realism.

The answer realism provides to the question under consideration does not have the difficulty of the answer provided by reductionism. The person who poses the question might not be satisfied with the answer because she believes that it does not state an accurate moral reason to do what is right. Indeed, this might be correct because the particular version of realism that provides the answer might be mistaken and therefore incapable of coming up with such a reason. However, the inquirer cannot protest that the answer merely repeats what she is asking about and therefore does not state the required reason. In contrast to the answer reductionism provides, the answer realism provides at least mentions a possible candidate for such a reason. Hence, this answer can, in contrast to the one reductionism offers, state a moral reason to do what is right.

This view of what a moral reason to do what is right consists in satisfies consequently the principle suggested above. Unlike reductionists, realists are not committed to the view that the property which constitutes a moral reason to do what is right consists in the property of being right. According to the example of a version of realism utilised above, the property that constitutes this reason is the property of maximising happiness, whereas one has this reason to do what is right. Since a moral reason to do what is right fulfils the mentioned principle, it might have normative force. The fact that it fulfils the mentioned principle underlines the view that realism is able to state a moral reason to do what is right.

As is clear from the reasoning above, it is thanks to the way realism reads the moral reason principle that it is able to give an adequate response to ‘What moral reason is there to do what is right?’ and hence maintain that there is a moral reason to do what is right. It is therefore reasonable to claim that realism offers a better interpretation of the moral reason principle than reductionism. That is, the moral reason entailed by the rightness of an action does not consist in the fact that the action is right, but in the fact that the action has a set of non-moral properties which is such that it makes actions right. The moral reason principle concerns the normativity of judgements to the effect that actions are right. Consequently, there is reason to assume that realism, unlike reductionism, provides a reasonable account of the normativity of such moral judgements.

## 6. *Three Objections*

There are at least three objections that might be raised against the line of argument of the previous section. As my responses to these objections are implied by what I have said earlier, I will make the discussion of them brief.

(i) Suppose it is asked ‘What moral reason is there to do what is right?’ Assume that the class of actions that is right might be picked out by means of some other property these actions have in common, the property A for example. It might then be objected that, instead of asking the mentioned question, it could just as well be asked ‘What moral reason is there to do what is A?’ It does not seem strange to answer this question ‘It’s right’. Hence, it might be thought that the above reasoning is mistaken.

As can easily be seen, this objection presumes the first interpretation of questions of the type ‘What reason is there to do F?’ considered in the last section. It illustrates thereby the importance of distinguishing between the two interpretations of this type of questions. Moreover, it illustrates how important it is for my arguments that the second interpretation, not the first one, is correct. Consequently, in the light of the discussion above, I think this objection is mistaken because it rests on the first interpretation. Contrary to what this interpretation suggests, in asking ‘What moral reason is there to do what is right?’, it is not merely asked what moral reason there is to perform a class of actions, a class of actions which might be picked out by means of some other of its properties than rightness. Rather, in line with the second interpretation, it is asked what moral reason there is to perform actions in special consideration of their having the particular property of being right. More exactly, it is asked what moral reason there is to perform action that are right irrespective of which other properties these actions might have (with the exceptions mentioned in the last section). Consequently, when it is asked ‘What moral reason is there to do what is A?’, quite another question is asked; it is asked what moral reason there is to perform actions that are A irrespective of which other properties these actions might have (with the mentioned exceptions). This question has quite another answer than the original one.

(ii) As we have seen, reductionism answers the question ‘What moral reason is there to do what is right?’ in terms of actions having a non-moral property which is claimed to be identical with rightness. A defender of the second objection points out that such an answer need not be felt to be odd

since the person who asks the question may not *know* that this non-moral property is identical with rightness.<sup>17</sup> Indeed, that the identity holds might not be known by anyone in which case the answer is not felt to be odd by anyone. And if the answer is not felt to be odd, the ground for rejecting it vanishes. Hence, it might be thought that the argument against the reductionist answer is mistaken.

In reply to this objection, it should first be pointed out that even if the answer reductionism gives is not experienced as strange because it is not known that rightness is identified with a certain non-moral property, it still does not provide a reasonable answer to the question. For, as I argued in the last section, the fact that an action has a non-moral property, where this property is identical with rightness, cannot constitute a moral reason to do what is right. Thus, even if the answer reductionism gives might *appear* to give such an answer, it does not. This contention is strengthened by the following consideration. Suppose someone asks ‘What moral reason is there to do what is right?’ The answer she gets is ‘It maximises happiness’. She may feel that this actually answers her question. However, later she gets to know that the person who answered the question takes rightness and maximising happiness to be identical. In that person’s vocabulary, to say that actions maximise happiness is just another way of saying that they are right. In light of this, the inquirer would presumably feel that she, after all, has not been given a satisfactory answer. She is likely to protest that if rightness and maximising happiness are thought to be the same, the answer to her question ‘What moral reason is there to do what is right?’ cannot be ‘It maximises happiness’. The answer she got just repeats what she asked about because on the view at issue, to answer ‘It maximises happiness’ is in effect the same as answering ‘It’s right’. Such a repetitious answer, she may argue, cannot provide the requested reason.

(iii) In the last section, I considered the question ‘What moral reason is there to do what is right?’, where the rightness referred to is moral rightness. According to the third objection, this question is inconceivable because it does not make sense to ask what *moral* reason there is to do what

---

<sup>17</sup> This objection is probably stronger when it is used by an adherent of synthetic reductionism than when used by an adherent of analytic reductionism. The reason is that on analytic reductionism, a person knows which non-moral property that constitutes the meaning of ‘right’ in so far as she is linguistically competent with respect to the meaning of the term. However, this is not the case according to synthetic reductionism.

is right, where rightness is understood to be *moral* rightness. In order for a question that asks what reason is to do what is morally right to make sense, 'reason' has to be understood as a reason that is not moral. Thus, it cannot be argued against reductionism that it is unable to provide an answer to this question.

There are at least three difficulties with this objection. First, in contrast to the two previous objections, I find it hard to think of any reason for it that does not beg the question. Of course, it might be claimed that it does not make sense to ask 'What moral reason is there to do what is right?', understood in the indicated way, because the fact that an action is right is identical with the pertinent moral reason. But, quite evidently, this argument begs the question against my reasoning. Second, as was implied above, there are reasons to believe that the question under consideration does make sense to ask. This is perhaps most clearly illustrated by the examples given above of various kinds of situations in which people may ask it. I cannot see but that the question makes perfectly good sense to ask in the kinds of situations in which these people find themselves. Third, and closely related to the last point, it is implausible to understand people who ask the question under consideration to be asking about another reason than a moral one. People who ask this question do not look at morality from outside and ask what reason there is to do what is right from such an external perspective. On the contrary, they find themselves in situations in which they are concerned about what reason morality provides to do what is right.<sup>18</sup> For instance, consider again the person in the first example above, a person who has started to mistrust morality. She realises that morality classifies certain actions as right but has come to suspect that the motivation for this is that of preventing the lower classes from revolting against the upper classes. However, she has not made up her mind yet as to whether her suspicion is correct. She then asks what reason morality in fact supplies to do what is right—what moral reason there actually is to do what is right. It is difficult to see that the question she is asking can be understood in any other way than as a question about a moral reason. Indeed, the very point of

---

<sup>18</sup> However, it should be recalled that according to some views of moral reasons, such reasons are identified with reasons belonging to some other normative perspective. Thus, suppose moral reasons are identified with rational reasons. In case the inquirer is aware of this, when she asks about a moral reason to what is right, she asks about a rational reason to what is right. But on the mentioned assumption, she would then not be asking about a reason to do what is right from a perspective external to morality.

her question is to ask what reason morality itself provides to do what is right. Surely, people may—and perhaps often do—ask about a reason to do what is right and have in mind a reason that is not moral. However, that is not the question they are asking when they ask the question under discussion here.

### *7. Analogous Normative Questions*

In section 5, I argued that the question ‘What moral reason is there to do what is right?’ is not satisfactorily answered by asserting ‘It’s right’, since, to put the point briefly, this answer merely repeats what it asked for and therefore cannot state the requested reason. Reductionism has to give this answer, as we saw, because it is committed to a certain reading of the moral reason principle. Accordingly, it might be argued that the way reductionism reads that principle is mistaken. This indicates in turn that it fails to provide a reasonable account of the normativity of judgements to the effect that actions are right. However, realism is not vulnerable to these objections.

It is interesting to note that a number of philosophers have argued in a similar manner. They have reflected on analogous normative questions and have found that these questions cannot be accurately answered in the repetitious manner in which reductionism answers the question considered above because such responses merely repeat what is asked for and therefore fails to state the requested reasons. Among these philosophers are John McDowell, George Nakhnikian, Philip Stratton-Lake, and, according to one interpretation, G. E. Moore.<sup>19</sup> It can be argued that what these authors maintain in relation to the questions they discuss lends confirmation to the

---

<sup>19</sup> McDowell (1978), p. 14; Nakhnikian (1963), pp. 157–158; Stratton-Lake (1999), pp. 80–83, and Stratton-Lake (2000), chap. 1. See also e.g. Dancy (2000), pp. 166–167. On the relevant interpretations of Moore, see Nakhnikian (1963), pp. 145–146, 156–158, and Regan (1972), pp. 48–58. Another author who has argued in a similar manner is R. M. Hare; see Hare (1952), chap. 5. Hare argues that if ‘good’ is defined in terms of a set of natural properties, the fact that something has these natural properties cannot be cited as a reason to get what is good; it cannot, in Hare’s terms, be used to commend what is thought to be good. However, Hare thinks that we sometimes want to commend things by appealing to properties that are employed in proposals of such definitions. It might be argued that Hare’s distinction between an evaluative term’s primary and secondary sense makes this possible. According to this proposal, the reason consists in the fact that something has the set of natural properties which constitutes the secondary sense of ‘good’, whereas this is a reason to have the attitude that constitutes the primary sense of the term. If this proposal is plausible, a version of non-cognitivism that employs the distinction between primary and secondary sense might be able to provide the requested answer to the question under consideration.

arguments put forward in section 5. Moreover, I think it can be argued that the explanation of why they find the mentioned kind of answer unsatisfactory corresponds to the explanation I proposed there.

Consider Stratton-Lake's discussion of the following question:

Why should I do what I believe I ought to do, rather than some other action? [- - -] [This question] asks why we should do what we believe we ought to do, rather than some other action. Suppose, for the sake of simplicity, we limit this question to cases in which our belief is true. [This question] may seem like a deep question if one thinks, as Kantians tend to, that the reason why one ought to  $\phi$  is because one ought to. My view is that this cannot be a normative reason why one ought to  $\phi$ . The normative reasons why I ought to  $\phi$  are those which would constitute a correct answer to the question 'why ought I to  $\phi$ ?'. But one cannot answer this question by replying, 'because you ought to  $\phi$ '. The fact that I ought to  $\phi$  cannot, therefore, be a reason why I ought to  $\phi$ . If this is correct, then [this question] may be easy to answer. The reason why you should do what you believe you ought to do may simply be the reason on the basis of which you believe you ought to do these various acts. Suppose you believe you ought to  $\phi$ , and that you believe this because you promised your friend that you would  $\phi$ . Why should you do what you believe you ought to do here? A naïve reply may be 'because you promised your friend that you would do this'.<sup>20</sup>

Stratton-Lake argues that the question 'Why should I do what I ought to do?' cannot be adequately answered by asserting 'Because you ought to', for such an answer fails to mention a normative reason.<sup>21</sup> The proper kind of answer to the question is instead given in terms of those features of actions on the basis of which they are such that they ought to be done.

The question Stratton-Lake focuses on runs, when formulated in terms of rightness: 'Why should I do what is right to do?' As we saw earlier, this question can be formulated thus: 'What reason is there to do what is right?' In asking the mentioned question, Stratton-Lake asks about a moral reason.<sup>22</sup> This is indicated by the context of the quotation, but also by the kind of answer he suggests: that the person in question has made a promise, a clearly moral consideration. Formulated in terms of rightness, the question Stratton-Lake poses becomes: 'What moral reason is there to do what is right?' Thus, he asks a question which is directly analogous to the one discussed in section 5. Moreover, Stratton-Lake's arguments in relation to this question are remarkably similar to the arguments I put forward there.

---

<sup>20</sup> Stratton-Lake (1999), pp. 82–83.

<sup>21</sup> Admittedly, Stratton-Lake starts with the question 'Why should I do what I believe I ought to do?' However, since he goes on to consider this question on the assumption that the inquirer's belief is true, it can be understood in the way I do.

<sup>22</sup> From the context it is also clear that Stratton-Lake takes 'ought' to refer to a moral ought.

First, he argues that the question he discusses cannot be accurately answered in the repetitious manner just mentioned because such an answer fails to state the requested reason. Second, he argues that the question instead is accurately answered in terms of actions having certain features on the basis of which they are such that they ought to be done. This idea is quite similar to the one proposed by realism concerning how the question discussed above should be answered, namely in terms of actions having a set of non-moral properties that makes them right.

Granted these similarities, it is reasonable to hypothesise that one underlying explanation of why Stratton-Lake finds that the question ‘Why should I do what I ought to do?’ is not satisfactorily answered by claiming ‘Because you ought to’ is analogous to the explanation I proposed above.<sup>23</sup> To be brief, the explanation would be this. The question Stratton-Lake poses, understood in a way corresponding to the second interpretation in section 5, is a fundamental normative question in relation to morality. A question of this type is not satisfactorily answered in the repetitious manner Stratton-Lake opposes because such an answer fails to state the requested reason. The underlying explanation can be provided in terms of the principle suggested above: a proposed reason to do what is F is such a reason only if the property that constitutes it is not identical with F. The reason suggested in the repetitious answer Stratton-Lake disapproves of does not fulfil this principle and, as a consequence, does not constitute the requested reason.

Consider briefly reductionism concerning ought. Following the reasoning in section 3, reductionism identifies the moral reason at issue with the fact that something has that moral feature. That is, reductionism is committed to reading an analogue to the moral reason principle in such a way that the moral reason at issue is identical with the fact that an action is such that it ought to be done. It is thus committed to answer the question Stratton-Lake poses in the repetitious manner he repudiates. According to the reasoning above, this indicates that reductionism is unable to account accurately for the normativity of judgements to the effect that actions ought to be done. By contrast, realism concerning ought is not committed to the

---

<sup>23</sup> In a later work, Stratton-Lake offers a slightly different explanation of why the question cannot be answered in the indicated manner: Stratton-Lake (2000), pp. 16–28. However, this explanation is not incompatible with the one I offer here; rather, they complement each other.



mentioned understanding of the relation between the fact that an action is such that it ought to be done and moral reason. It is therefore not committed to giving the repetitious answer to the question posed by Stratton-Lake, something which suggests that it is capable of accounting for the normativity of such moral judgements. As already indicated, I think the same kind of reasoning is applicable to arguments other authors have proposed in relation to analogous normative questions.<sup>24</sup>

---

<sup>24</sup> One author who propounds a similar argument is George Nakhnikian. What he says is especially relevant to one of the objections I considered in the last section, objection (ii). There I imagined that it might be objected that the answer reductionists give to the question ‘What moral reason is there to do what is right?’, may not be felt as awkward because the inquirer might not be aware of the identification advanced by the reductionist who gives the answer. It might then be thought that the argument against reductionism is mistaken. It is noteworthy that Nakhnikian answers a related objection in virtually the same way as I do. He writes: ‘Consider this example. I am trying to show you that doing A is your duty. You are not convinced. But A is your duty, I say, because you doing A would contribute to the welfare of the community. But suppose that I define “A is your duty” to mean that your doing A will contribute to the welfare of the community. In that case, “A is your duty” is just another way of saying “Your doing A will contribute to the welfare of the community,” and I have not provided a reason why doing A is your duty. In asserting “A is your duty” and “Your doing A will serve the welfare of the community” I have simply used two different expressions to say the same thing. This is illuminating only to those who need a lesson in language. It is of no help whatever to anyone who requires to be shown that doing A is his duty’ (Nakhnikian (1963), pp. 157–158). Translated to the question I focus on, Nakhnikian’s point can be formulated thus: on the assumption that rightness is identical with a given non-moral property, such as contributing to the welfare of community, the question ‘What moral reason is there to do what is right?’ is not satisfactorily answered by claiming ‘It contributes to the welfare of community’, even if the person who asks the question is not aware of the identification. The argument Nakhnikian suggests for why this is so corresponds to the one I offered above. It is, roughly put, that even if it is not apparent to the one who asks the question, such an answer is just another way to express what is asked about and it fails therefore to offer the requested reason.

It might perhaps also be argued that the reasoning in section 5 applies to certain normative questions concerning moral issues, although they are not directly analogous to the type of questions we have so far discussed. Consider, for example, the question ‘Why should I be moral?’ As noticed earlier, questions that ask why something should be so and so ask about a reason. Moreover, some of these questions ask about a moral reason. The phrase ‘to be moral’ might be taken to refer to a disposition to perform actions that are right. Thus understood, the question asks what moral reason there is—what reason morality provides—to have the disposition to do what is right. Like the question considered above, this question seems intelligible to ask. The discussion above about the moral reason principle indicates that reductionism has to answer this question by claiming, in effect, that the moral reason to be moral—to have the disposition to do what is right—is that it is right. This answer is flawed for the same reason as the answer reductionists give to the question considered above is: it merely repeats what is asked for and does consequently not state the requested reason. The underlying explanation of why this response is unsatisfactory can also be claimed to correspond to the explanation above. On the realist reading of the moral reason principle, the answer is given in terms of actions having a set of non-moral properties which is such that it makes actions right. According to the version of realism appealed to earlier, the answer is that a moral reason

Moreover, it seems plausible to argue that my reasoning puts some light on what T. M. Scanlon refers to as ‘Prichard’s dilemma’. Scanlon writes:

Attempts to explain how the fact that an action is wrong provides a reason not to do it face a difficult dilemma. Understood in one way, the answer is obvious: the reason not to do the action is just that it is wrong. But this is surely not the kind of answer that is wanted: it simply takes the reason-giving force of moral considerations for granted. Suppose, on the other hand, that we were to appeal to some clearly nonmoral reason, such as that people have reason to be morally good because, taking into account the effort that deception requires, the likelihood of being found out, and the costs of social ostracism, it is in their self-interest to be moral. This account might supply a reason for doing the right thing, but it would not be the kind of reason that we suppose a moral person first and foremost to be moved by. I will refer to this as Prichard’s dilemma. So a satisfactory answer to our question must not, on the one hand, merely say that the fact that an action is wrong is a reason not to do it; but it must, on the other hand, provide an account of the reason not to do it that we can see to be intimately connected with what it is to be wrong. Answers can thus be arranged along one dimension according to their evident moral content, ranging from those that appeal to what seem most obviously to be moral considerations (thus running the risk of triviality) to those having the least connection moral notions (thus running the risk of seeming to offer implausibly external incentives for being moral).<sup>25</sup>

Scanlon is concerned with the following question: ‘What reason is there not to do what is wrong?’ From the context of the quotation, it is clear that the wrongness at issue is moral wrongness. Scanlon suggests that the kind of reason referred to in this question is ‘the kind of reason that we suppose a moral person first and foremost to be moved by’, and it seems therefore not far-fetched to assume that he has in mind a moral reason.<sup>26</sup> Hence, it seems reasonable to assume that the question Scanlon poses is directly analogous to the fundamental normative questions about moral issues discussed earlier. Scanlon considers two answers to the question. The first answer is ‘It’s wrong’. Scanlon finds this answer unsatisfactory because ‘it takes the reason-giving force of moral considerations for granted’ and is ‘running the risk of triviality’. The second answer is ‘It’s in your self-interest’. Scanlon finds this answer unsatisfactory because it does not state the kind of reason asked for, i.e. a moral reason. The fact that both answers are inadequate gives rise to what Scanlon calls ‘Prichard’s dilemma’. The dilemma demonstrates,

---

to be moral—to have the disposition to do what is right—is that this maximises happiness. Such an answer may state the reason asked for.

<sup>25</sup> Scanlon (1998), pp. 149–150.

<sup>26</sup> This is also indicated by the fact that Scanlon regards it to be uncontroversial that if an action is wrong, there is a reason not to do it. Moreover, it is indicated by the fact that he contrasts the kind of reason he seeks with a ‘nonmoral reason’.

Scanlon believes, that a satisfactory answer to the question at issue cannot just be ‘It’s wrong’, but at the same time it has to mention a reason that is ‘intimately connected to what it is to be wrong’.<sup>27</sup> Prichard’s dilemma deserves detailed discussion. Here I shall only make three comments. First, as should be familiar by now, the reasoning above can explain why ‘It’s wrong’ is an unsatisfactory answer to the question Scanlon poses. Second, it can also explain why he finds this answer unsatisfactory. To answer ‘It’s wrong’ can be said to take the reason-giving force of moral considerations for granted because it merely presumes that if an action is wrong, it follows that there *is* a reason not to do it, but it does not state *what* this reason consists in. The answer is also, in a certain sense, trivial since it only repeats what is asked about. Third, it is plausible to argue that according to the realist conception of moral reasons, there is a moral reason that fulfils Scanlon’s requirement on a satisfactory answer to the question he poses. On this conception, the moral reason at issue consists in the fact that an action has a set of non-moral properties which is such that it makes actions wrong. Since this set of non-moral properties is not identical with wrongness, the realist answer to Scanlon’s question is not merely ‘It’s wrong’. But since such a set of non-moral properties *makes* actions wrong, this answer certainly states a reason that is ‘intimately connected to what it is to be wrong’.<sup>28</sup>

---

<sup>27</sup> Scanlon refers to the discussion in Prichard (1952 (1912)), pp. 149–162. However, it is not obvious that Prichard is concerned with the same dilemma that Scanlon formulates. For an interpretation and critical discussion of Prichard’s essay, see Schwarz (1971), pp. 169–180.

<sup>28</sup> Scanlon argues that ‘an action is wrong if its performance [. . .] would be disallowed by any set of principles [. . .] that no one could reasonably reject as a basis for informed, unforced general agreement’ (Scanlon (1998), p. 153). He seems to want to characterise the moral reason not to do what is wrong in these terms (see e.g. Scanlon (1998), p. 155). It is controversial whether it is best to construe this view in such a way that it reduces, and hence *identifies*, wrongness with being ‘disallowed by any set of principles that one could reasonably reject...’ or in such a way that it says that what *makes* actions wrong is the fact that they have this feature (where ‘make’ is understood as non-reductive relation). It might be argued that what I said above lends support to the latter alternative. (However, there are indications that Scanlon himself prefers the first alternative; see Scanlon (1998), pp. 10–12.) Suppose that the first alternative is opted for. In that case, it is difficult to see that Scanlon can avoid the first horn of the dilemma he presents. According to this alternative, wrongness is identified with being ‘disallowed by any set of principles that one could reasonably reject...’ and the pertinent moral reason is also understood as the fact that an action has that feature. As a consequence, the moral reason is identical with the fact that an action is wrong. Hence, the answer to the question Scanlon poses is in effect ‘It’s wrong’. On this alternative, Scanlon’s own view would be exposed to the first horn of the dilemma. However, if the second alternative is opted for, the first horn of the dilemma is avoided. According to this alternative, the

## 8. *Other Normative Questions*

In the previous sections, I maintained that reductionism cannot give satisfactory answers to certain fundamental normative questions about moral issues, but that realism is able to do so. However, it should be observed that reductionism seems to deliver acceptable answers to many familiar normative questions about moral issues. Consequently, reflections on these questions do not help us to decide which view to prefer. To see this, we may recall some of the normative questions about moral issues mentioned in section 5. Some such questions, we noticed, ask what moral reason there is to perform a particular action. We might for example ask ‘Why should I give him the money back?’ and take it as a question about a moral reason. That is, we might ask ‘What moral reason is there to give him the money back?’ Return to our example of reductionism utilised earlier. According to this view, the answer to the question is ‘It maximises happiness’. Since rightness and maximising happiness are identical on this view, the answer is in effect ‘It’s right’. I do not think either version of the answer reductionism provides strikes us as clearly mistaken. Realism answers the question in terms of a set of non-moral properties that makes actions right but with which rightness is not identified. According to the example of a version of realism utilised above, the answer would be ‘It maximises happiness’. I do not think this answer strikes us as clearly preferable to the answer offered by reductionism.

It should be pointed out, however, that the fact that the answers realism gives to certain normative questions do not seem to be superior to the ones reductionism gives does not show that the arguments of the previous sections are flawed. What it shows is rather that the advantages realism has over reductionism with regard to reason become manifest only when we consider the answers these two views give to certain particular kinds of normative questions. As we have seen, these are fundamental normative questions about moral issues which ask what reason morality itself

---

moral reason is characterised in terms of being ‘disallowed by any set of principles that one could reasonably reject...’. However, wrongness is not identical with this feature. This means that the question Scanlon poses can be answered in these terms without the first horn of the dilemma emerging. For relevant discussions of how Scanlon’s view should be construed and matters related to this, see McNaughton and Rawling (2003), pp. 328–331; Parfit (2003), pp. 389–390, n. 21; Pettit (2000), pp. 156–163; Sosa (2004), pp. 374–375, and Stratton-Lake (2003), pp. 70–76.

provides to do what has a certain moral feature. That is, they are questions that utilise moral notions both in the reason asked for and in that for which a reason is asked.

The fact that reductionism seems to give acceptable answers to some normative questions may however help to explain why it might not be recognised that reductionism fails in the ways I have maintained. The explanation is simply that the focus has been on the familiar and common kind of normative questions just exemplified rather than on the fundamental kind of normative questions discussed earlier.

A further explanation is worth mentioning. There are normative questions that are similar to the one considered in section 5 but which do not make sense. Take the question ‘Do I have a moral reason to do what is right?’ This question does not make sense, since if an action is right it follows, according to the moral reason principle, with analytic necessity that there is a moral reason to perform it. Earlier we considered a version of the question ‘Why should I do what is right?’ Assuming that it asks about a moral reason, this question is most reasonable interpreted as ‘What moral reason is there to do what is right?’ In asking this question, one presupposes that there *is* a moral reason to do what is right; one asks *what* such a reason consists in.<sup>29</sup> As we have seen, there is reason to believe that such a question makes sense. However, the question ‘Why should I do what is right?’ might perhaps also be interpreted similarly to the nonsensical question just mentioned. In asking the question thus understood, one asks *whether* there is a moral reason to do what is right.<sup>30</sup> Since this question does not make sense, it cannot be argued against reductionism that it fails to provide an answer to it. If the fundamental normative questions about moral issues considered above have been interpreted as questions like this one, this might help to explain why it has not been recognised that reductionism fails in the ways I have maintained.

### 9. *A Note on Rationalism and Realism*

Above I have maintained that the reading of the moral reason principle provided by realism is superior to the reading provided by reductionism. The difference in their reading of this principle has, I maintained,

---

<sup>29</sup> Cf. Nielsen (1984), pp. 81–82.

<sup>30</sup> Cf. Frankena (1976 (1958)), p. 66.

consequences for the plausibility of reductionism and realism, respectively. In spite of these results, the arguments offered with regard to the moral reason principle are purely formal, since they merely concern how the principle should be read internally, or, more precisely, how the relation between the fact that an action is right and moral reason is to be construed. Thus, the arguments do not presuppose that moral reasons are understood in any particular way. This is as it should be since it is desirable that a reading of the moral reason principle is independent of various views of moral reasons in order not to beg significant questions. One of the most discussed issues in meta-ethics is rationalism.<sup>31</sup> It is relevant to consider whether the way realism reads the moral reason principle is compatible with this view. If it is not, it might be argued that it is not neutral in the desirable way. In section 2, I mentioned that rationalism can be understood as the view that moral reasons consist in reasons of rationality, what I have referred to as ‘rational reasons’. As mentioned before, we may formulate this view in the following way:

*Rationalism:* It is necessary that if an action is right, then there is a rational reason to perform that action.<sup>32</sup>

In section 4, we saw that on the realist reading of the moral reason principle, a moral reason consists in the fact that an action has a set of non-moral properties which is such that it makes actions right. Rationalism can, as just mentioned, be understood as the view that moral reasons consist in rational reasons. According to a rationalist version of realism, a rational reason of the relevant kind consists consequently in such a fact. Hence, the way realism understands the moral reason principle seems able to accommodate rationalism in a relatively straightforward way.

It might be argued, however, that even if the realist reading of the moral reason principle is compatible with rationalism understood in a general way, it does not show that it is compatible with rationalism in combination with certain views about what rational reasons consist in. This is a large issue that I cannot deal with here. However, I would like very briefly to consider a conception of rational reasons that, in combination

---

<sup>31</sup> The same question can in principle be asked in relation to some other kind of reasons. However, since the relation moral judgements have to rational reasons presumably is considered most central, the issue discussed here seems to be the most important one.

<sup>32</sup> Cf. Brink (1992), pp. 1–2, 8–9; Smith (1994), p. 62, and Shafer-Landau (2003), p. 170.

with rationalism, might be thought to constitute a problem for meta-ethical views known as 'realism'. On this conception of rational reasons, having such a reason to perform an action is a function of whether this would contribute to the satisfaction of some of one's desires. Of course, this is merely a very rough formulation of a type of view which exists in various elaborated versions.<sup>33</sup> Rationalism states that, necessarily, if an action is right, there is a rational reason to perform that action. The result of the combination of these views is that the rightness of an action is dependent on whether its performance would satisfy the agent's desires.<sup>34</sup> Is this version of rationalism compatible with the way realism reads the moral reason principle? I think it is, at least if realists understand it in a certain way. There seems to be a possible version of realism which says that what makes an action right is a set of non-moral properties which pertains to desires in the mentioned way. Thus, the way realism reads the moral reason principle seems compatible with rationalism in combination with the mentioned view of rational reasons. It remains, though, to account for how such a version of realism should be understood.<sup>35</sup>

However, even if the way realism reads the moral reason principle is compatible with the combination of rationalism and the mentioned conception of rational reasons, there are indications that the resulting view should be rejected. One of our fundamental notions of morality is that requirements of moral rightness are categorical in the sense that they apply to people independently of what they happen to desire. As the version of realism under consideration appears to be incompatible with this notion, there is reason to believe that it should be rejected.<sup>36</sup> There are basically two ways to do so. Either rationalism might be denied or this conception of

---

<sup>33</sup> The most well-known version is presumably the one put forward in Williams (1981 (1980)), pp. 101–113.

<sup>34</sup> This is of course a much-discussed issue in moral philosophy. For helpful discussions of this and related issues, see e.g. Brink (1989), pp. 50–78; Brink (1992), pp. 1–26; Cullity and Gaut (1997), pp. 1–23, and Shafer-Landau (2003), chap. 7 and 8.

<sup>35</sup> There is at least one more version of rationalism that might be difficult to combine with realism. According to this version, rationalism entails internalism as regards moral motivation; see Smith (1994), p. 62. (For criticism of this version of rationalism, see e.g. Miller (2003), pp. 229–232.) The reason why this view might be thought to be problematic for realism is that it might be argued that internalism is incompatible with realism. However, in the next chapter, I will argue that internalism is mistaken.

<sup>36</sup> However, it should be mentioned that there are theories of rational reasons that might be seen as versions of the view mentioned above but that are claimed to be compatible with the notion that requirements of moral rightness are categorical; see e.g. Smith (1994), pp. 164–175.

rational reasons might be denied.<sup>37</sup> In fact, I think both these views should be denied. However, this is not something I will argue for here.

### **10. *Concluding Remarks***

According to the moral reason principle, moral judgements to the effect that actions are right are normative in the fundamental sense that it is analytically necessary that if an action is right, there is a moral reason to perform it. Reductionism reads this principle in such a way that the moral reason is identified with the fact that the action is right. However, realism may read this principle in such a way that the moral reason consists in the fact that the action has a set of non-moral properties which is such that it makes actions right. I have argued that because of this difference, reductionism fails to answer a version of the question ‘Why should I do what is right?’, whereas realism is capable of doing so. Moreover, I have argued that this indicates that reductionism fails to provide an accurate account of the normativity of the mentioned kind of moral judgements, whereas realism is capable of doing so. I have also suggested that similar lines of reasoning apply to analogous normative questions concerning moral issues. I finished the chapter by suggesting that the way realism understands the moral reason principle is compatible with rationalism. However, this issue requires further discussion.

---

<sup>37</sup> For various strategies how this might be done, see e.g. Brink (1992), pp. 7–14.



## Chapter 7

### Explaining Moral Motivation

#### 1. *Introduction*

In the last chapter, I mentioned that there are basically two ways in which moral judgements can be claimed to involve reasons for actions: by involving normative reasons and by involving motivating reasons. I devoted that chapter to a discussion of the relation between moral judgements and normative reasons of a certain kind, viz. moral reasons. The present chapter is devoted to a discussion of the relation between moral judgements and motivating reasons.

One of the most debated views in meta-ethics is internalism as regards moral motivation. According to a well-known argument which I will refer to as ‘the internalist argument’, internalism in conjunction with the Humean theory of motivation implies that moral judgements do not consist in beliefs. The internalist argument means consequently that cognitivism is false and provides support to non-cognitivism. In the present chapter, I will argue that the internalist argument should be rejected and that cognitivism therefore can be maintained. I will do so by arguing against internalism and in favour of its opposite, externalism. In order for this argument to be plausible, advocates of externalism have to be able to account for moral motivation. As a part of the argument, I will maintain that realism, unlike reductionism, succeeds to account for how a pertinent aspect of moral motivation should be understood on externalism. Consequently, realism, unlike reductionism, plays a part in the defence of cognitivism against the internalist argument. However, in comparison with the previous two chapters, the contribution of realism to the meta-ethical issue dealt with here is considerably more limited.

In the next section, I formulate internalism and externalism, and in the section thereafter I formulate the internalist argument. In section 4, I examine the most important considerations that are appealed to in defence of internalism and argue that they do not provide such a strong support for this view that one might assume. In section 5, I start a discussion about a possible counterexample to internalism, the renowned amoralist, and the

different ways internalists have argued that there are no amoralists. In section 6, I continue this discussion by arguing that there are descriptions of amoralists—what I will refer to as ‘advanced amoralists’—that both are compatible with the internalist considerations and cannot be dismissed in the ways proposed by internalists. In section 7, I argue by means of some ideas in pragmatics that externalism can explain the considerations that are thought to provide support to internalism and that these explanations are preferable to those offered by internalism. The most interesting argument for internalism and against externalism is Michael Smith’s so-called fetishist argument. I have saved this argument to a separate discussion in section 8 where I argue that it is unsuccessful. Smith’s argument raises the question how moral motivation should be understood according to externalism. In section 9, I argue that realism provides a better account than reductionism of a certain type of motivational states which is important to the plausibility of externalism.

## 2. *Internalism and Externalism*

As far as I understand, it is generally agreed in meta-ethics that there is a significant connection between moral judgements and motivation that calls for explanation. However, there are different views about how this connection should be accounted for; in particular, there is much disagreement about how strong it is. According to an influential view, this connection is analytically necessary. Applied to moral judgements to the effect that actions are right, this view yields the following principle:<sup>1</sup>

---

<sup>1</sup> There are a number of various theses that are labelled ‘internalism’, including what I referred to as rationalism in the previous chapter. In this chapter, I will only be concerned with internalism as regards moral motivation. For overviews of various types of internalism, see e.g. Audi (1997), pp. 125–159; Brink (1989), pp. 37–43; Parfit (1997), pp. 99–105, and Shafer-Landau (2003), pp. 142–145. For useful discussions about the definition of internalism as regards moral motivation, see e.g. Cuneo (1999), pp. 361–363; Korsgaard (1996 (1986)), pp. 315–316, and Svavarsdóttir (1999), pp. 163–165. This kind of internalism is embraced by a number of authors, and it might be argued that this view has deep historical roots; for arguments to this effect, see e.g. Strandberg (2000), pp. 71–87. For some instructive statements of this kind of internalism, see e.g. Blackburn (1984), p. 188; Falk (1952 (1948)), p. 494; Hare (1952), pp. 20, 168–169; Korsgaard (1996 (1986)), pp. 311–334; Smith (1994), chap. 3, and Stevenson (1963 (1937)), p. 13. However, not all these authors understand internalism as regards moral motivation in the way I do.

*Motivational internalism:* It is analytically necessary that if a person judges that it is right for her to perform an action, then she is at least somewhat motivated to perform that action.<sup>2</sup>

According to motivational internalism—‘internalism’ for short—it is analytically necessary that if a person holds a judgement to the effect that an action is right, she is motivated to act in accordance with it. As I will understand internalism, it states that a person’s motivation to do what she judges to be right is *internal* to her judgement. There are, as far as I see, primarily two ways in which motivation can be internal to a person’s moral judgement: either in so far as the judgement consists in a motivational state of the mentioned kind or in so far as the judgement by itself gives rise to such a motivational state, presumably by causing it.<sup>3</sup> That the motivation is internal to a person’s moral judgement implies that her judgement is sufficient for her to be motivated to perform the action. There is thus no need for her to be in any state that is not internal to her judgement in this way in order to be motivated to perform the action; that is, there is no need for her to be in a motivational state that is *external* to her judgement to be so motivated. In the next section, I will consider an important qualification of internalism.

It is generally presumed in discussions about internalism that this view states that motivation is internal to a person’s moral judgement in the way described, although it is not always spelled out. This assumption is often put by saying that, according to internalism, moral judgements necessarily motivate or that they are necessarily motivating. However, it should be noted that if internalism is not understood to say that motivation is internal to a person’s moral judgement in the indicated way, it would not be able to function as a premise in the internalist argument. I will return to this consideration in the next section.

Since internalism is claimed to be analytically necessary, a person does not really judge that an action is right unless she is motivated to perform it

---

<sup>2</sup> As I mentioned in chapter 1, I use ‘judgement’ and ‘judge’ in a way that is neutral between cognitivism and non-cognitivism. If cognitivism is correct, a moral judgement consists (primarily) in a belief; if non-cognitivism is correct, it consists (primarily) in a non-cognitive state. In writing about internalism below, I will omit the phrase that the person in question judges that the action is right ‘for her’ to perform. However, I will take it to be implied in what follows.

<sup>3</sup> There might be variants of these two basic possibilities. According to one such variant, a person’s moral judgement partly consists in a motivational state and so entails it. I will take such variants to be covered by the formulations mentioned in text.

by a motivational state that is internal to her judgement. Similarly, to know that ‘right’ is to be applied in accordance with the principle is a condition for being linguistically competent with respect to the term.

Motivational externalism—‘externalism’ for short—is the denial of internalism.<sup>4</sup> Thus, externalism denies that it is analytically necessary that if a person holds a judgement to the effect that an action is right, she is motivated to act in accordance with it. As I will understand externalism, it states that motivation is not internal to a person’s moral judgement in the way proposed by internalism. That is, her judgement neither consists in a motivational state nor gives rise to it by itself. That motivation is not internal to a person’s moral judgement means that her judgement is not sufficient for her to be motivated to do what she judges to be right. In order for her to be motivated to do what she judges to be right, she has to be in an external motivational state; that is, she has to be motivated by a state that is not internal to her judgement. External motivational states may consist in states that involve moral concerns, but they may also, for example, consist in states that involve merely egoistic concerns.

According to externalism, a person may genuinely judge that an action is right without being motivated to perform it; in particular, she does not need be motivated by a motivational state that is internal to her judgement. Moreover, taking ‘right’ to apply in accordance with the internalist principle is not a condition for being linguistically competent with respect to the term.

The difference between internalism and externalism can be stated in terms of the distinction between normative and motivating reasons. Internalists are committed to the view that, with analytic necessity, if a person judges that an action is right, she has a motivating reason to perform that action. Externalists are committed to the denial of this view. However, internalists and externalists are not committed to any particular view of normative reasons.<sup>5</sup> In order not to confuse normative with motivating reasons, I will not make use of the term ‘reason’ in this chapter.

---

<sup>4</sup> Externalism is defended by e.g. Brink (1989), pp. 37–50; Copp (1995), pp. 187–219; Shafer-Landau (2003), chap. 6; Svavarsdóttir (1999), pp. 161–219, and Zangwill (2003), pp. 143–154.

<sup>5</sup> Cf. Smith (1994), p. 96.

### 3. *The Internalist Argument*

Internalism is widely considered to have far-reaching implications for the correctness of meta-ethical views. In particular, according to a very influential argument, internalism together with what is known as the Humean theory of motivation implies that cognitivism is false. The basic claim of the Humean theory of motivation is that beliefs are not sufficient for motivation. Thus, the theory says, in Terence D. Cuneo's formulation, that 'beliefs are not identical with, nor do they generate by themselves, desires or motivational states'.<sup>6</sup> In addition, the theory claims that non-cognitive states, notably desires, are necessary for motivation to occur. (However, the latter claim is not a necessary part of the argument.)

The argument I have in mind—the internalist argument—can be represented as follows:<sup>7</sup>

- (1) *Motivational internalism*: It is analytically necessary that if a person judges that it is right for her to perform an action, then she is at least somewhat motivated to perform that action.
- (2) *The Humean theory of motivation*: Beliefs are not sufficient for motivation; beliefs do not consist in motivational states, nor do they give rise to motivational states by themselves.
- (3) Therefore, a person's judgement to the effect that an action is right for her to perform does not consist in a belief.

We can now see why it is essential for the internalist argument that internalism is understood to say that a person's motivation to do what she judges to be right is internal to her judgement in such a way that her judgement is sufficient for her to have that motivation. Premise (1), internalism, together with premise (2)—the view that beliefs are not sufficient for motivation, neither by consisting in motivational states nor by giving rise to such states by themselves—should yield the conclusion (3), that a person's judgement to the effect that an action is right does not

---

<sup>6</sup> Cuneo (2002), p. 467. It might be argued, however, that this common formulation of the Humean theory of motivation is not in accordance with Hume's own view; see Persson (1997), pp. 211–228.

<sup>7</sup> For clear formulations of the argument, see e.g. Brink (1997), p. 6; Parfit (1997), p. 106, and Shafer-Landau (2003), p. 121. In the literature, the argument more often seems assumed than clearly stated; but see e.g. Blackburn (1984), pp. 188–189. However, it is a matter of interpretation how proponents of this line of reasoning would exactly formulate the internalist argument.

consist in a belief.<sup>8</sup> This means that for (3) to follow, internalism has to be understood to state that the person's judgement *is* sufficient for motivation, either in so far as it consists in a motivational state or in so far as it gives rise to such a state by itself.

The conclusion (3) entails that cognitivism as regards judgements to the effect that actions are right is false.<sup>9</sup> (However, it might be argued that there are analogues to the internalist argument that pertain to other kinds of moral judgements.) Moreover, the argument provides support to non-cognitivism.

Cognitivism can be defended against the internalist argument in basically two ways. Which line of defence cognitivists adopt depends on whether they want to save internalism, the Humean theory of motivation or neither. First, cognitivists who are internalists dispute premise (2), the Humean theory of motivation.<sup>10</sup> According to one version of this option, some beliefs, among which are found not only moral but also non-moral beliefs, are sufficient for motivation. According to another version, only moral beliefs are sufficient for motivation whereas other kinds of beliefs are not. Proponents of this line of thought also differ as to whether the beliefs in question are motivating by themselves or because they generate, presumably by causing, motivating states. Second, cognitivists who are externalists dispute (1). This is also the line of defence that cognitivists who want to defend the Humean theory of motivation have to adopt. Cognitivist externalists who are not committed to the Humean theory of motivation are of course also free to reject (2).

In this chapter, I will defend cognitivism by challenging premise (1), internalism, and defend externalism in combination with realism. To my

---

<sup>8</sup> As we saw in chapter 1, according to some versions of non-cognitivism, moral sentences do express beliefs, although secondarily. The internalist argument above may thus need to be reformulated to take this into account.

<sup>9</sup> Since beliefs are expressed in sentences that have truth-value, this conclusion can also be formulated by saying that moral sentences to the effect that actions are right do not have truth-value.

<sup>10</sup> See e.g. Dancy (1993), esp. chap. 1; McDowell (1978), pp. 13–29; McDowell (1979), pp. 331–350; McNaughton (1988) pp. 108–113; Nagel (1970), chap. 1, and Wiggins (1991), 51–85. For critical views on some aspects of this view, see Strandberg (1999), pp. 171–192. At least some authors in this tradition seem to have in mind another notion of motivational internalism than the one I formulated above. In particular, some of these authors seem to believe that moral judgements, i.e. moral beliefs, need not always be sufficient for motivation, merely that they *sometimes* are sufficient for motivation; see e.g. McNaughton (1988), pp. 134–135.

mind it might be the case that premise (2), the Humean theory of motivation, is mistaken, but this is not something for which I will argue.

Before we proceed, I would like to comment on the formulation of internalism. Some authors defend a weaker claim than the one I have given this label. Michael Smith advocates what he calls ‘the practicality requirement’: ‘If an agent judges that it is right for her to  $\phi$  in circumstances C, then either she is motivated to  $\phi$  in C or she is *practically irrational*.’<sup>11</sup> Unfortunately, Smith does not explain what practical rationality amounts to, but he suggests that it excludes ‘distorting influences of weakness of will and other similar forms of practical unreason’.<sup>12</sup> Smith’s reason to incorporate a requirement of practical rationality into internalism is that it is ‘manifestly implausible’ to deny that ‘weakness of the will and the like may defeat an agent’s moral motivations while leaving her appreciation of her moral reasons intact’.<sup>13</sup>

There are basically two reasons why I hesitate to incorporate a requirement of practical rationality into internalism.

First, it would run the risk of committing internalists to a particular view of judgements to the effect that actions are right. If a requirement of practical rationality is incorporated into internalism, it has to be something about such judgements that explains why it is only if one is practically rational that one is guaranteed to be motivated in accordance with them. The explanation that most readily comes to mind is that such judgements pertain to rational considerations in some way, e.g. by being about what people would be motivated to do were they rational.<sup>14</sup> If this is correct, the principle would commit internalists to a certain understanding of these judgements. However, this is undesirable because it might not be a view that internalists are prepared to accept. Moreover, it is undesirable to make internalism conditional on a view that presumably is more controversial than internalism itself.

---

<sup>11</sup> Smith (1994), p. 61. Italics added. Cf. Korsgaard (1996 (1986)), pp. 315–317.

<sup>12</sup> Smith (1994), p. 61. Smith’s concept of practical rationality should not be conflated with his concept of full rationality which occurs in his analysis of ‘right’; although he does not explain the first concept, he gives an elaborated explication of the latter; see Smith (1994), pp. 155–161.

<sup>13</sup> Smith (1994), p. 61. Michael Stocker argues in a well-known paper that there are cases where people’s motivation does not accord with their evaluative judgements because of depression, apathy and the like; see Stocker (1979), pp. 738–753.

<sup>14</sup> Of course, this is Smith’s view.

Second, it would run the risk of getting internalists into a dilemma. On the one hand, if the notion of practical rationality is strong, so that people have to be practically rational in a quite substantial way to be guaranteed to be accordingly motivated, this seems to put considerable limits on how extensive the correlation between utterances of moral sentences and motivation can be. In that case, there is reason to suspect that internalism cannot explain the close correlation between them that often it is thought to hold. This would be especially troublesome since one argument in favour of internalism is considered to be that it is able to explain this correlation. On the other hand, if the notion of practical rationality is weak, so that little or no substantial demands are put on being practically rational, there is reason to suspect that it does not exclude the states of ‘practical unreason’ Smith wants it to exclude. In that case, internalism would not be able to account for the exceptions to the connection between moral judgements and motivation in the way Smith wishes. Hence, it would not be able to avoid a view Smith believes is ‘manifestly implausible’. Unfortunately, Smith does not give us an account of practical rationality that can help us to settle whether this dilemma can be avoided.<sup>15</sup>

We can now see that a commentator on the internalist argument finds herself in another dilemma. On the one hand, if a requirement of practical rationality is incorporated into internalism, the problems just mentioned emerge. On the other hand, if internalism is understood in the stronger way without a requirement of practical rationality, internalism is not able to avoid a view Smith thinks is ‘manifestly implausible’. On either alternative, the internalist argument seems to have lost in strength. The most satisfactory solution would be to propose a formulation of internalism that avoids this dilemma. Unfortunately, I do not know of such a formulation.<sup>16</sup> I have therefore chosen the following strategy. In the discussion below, I will in general continue to understand internalism in the stronger way described in the last section. However, in my arguments against internalism, I will not appeal to states that qualify as practical irrationality. This approach might not be entirely satisfactory but it avoids the dilemma. It avoids the relevant parts of the first horn of the dilemma for the following reasons. Understood in

---

<sup>15</sup> For a related problem with the practicality requirement, see Miller (2003), p. 221.

<sup>16</sup> Cf. Svavarsdóttir (1999), pp. 164–166.



this way, internalism does not appeal to a notion of practical rationality that the pertinent judgements have to comprise and there is thus no reason to believe that it commits internalists to a particular understanding of these judgements in the way mentioned above. Neither is there any immediate reason to believe that it is not able to explain the tight correlation that is assumed to hold between utterances of moral sentences and motivation. It avoids the second horn of the dilemma for the following reason. Since I do not appeal to states that qualify as practical irrationality in my arguments against internalism, I leave open the possibility that there is a formulation of internalism that accounts for why there might be exceptions to the assumed connection between moral judgements and motivation because of practical irrationality. Hence, I do not appeal to a view that is, as Smith says, 'manifestly implausible'.<sup>17</sup> In understanding internalism in this way, I admit for the possibility that there is a version of this view that steers clear of the dilemma. One way internalism might avoid the dilemma is if there is a formulation of 'practically rational' that does not have the mentioned problems. Since there might be such a formulation, I will at the relevant points in the discussion assume that advocates of internalism can appeal to a formulation of that kind in order to avoid arguments directed against this view.

#### 4. *Internalist Considerations*

In the present section, I will examine the considerations that, in my view, constitute the most significant support to internalism. It is customarily taken for granted, also by those who deny internalism, that these considerations make this view seem plausible. However, I will suggest that on closer examination these considerations do not have the strength that they often are assumed to have. As it is not possible for me to discuss all considerations that have been thought to support internalism, I will limit my discussion to the three that, to my mind, are most central.<sup>18</sup>

---

<sup>17</sup> For the same reason, this strategy avoids the second horn of the dilemma that internalists run the risk of getting into.

<sup>18</sup> There are a number of traditional arguments for internalism that I will not discuss, among others R. M. Hare's argument from practicality. The main reason why I do not discuss these arguments is that few moral philosophers, internalists included, nowadays seem to believe that they provide support to internalism. These arguments are consequently seldom considered in recent literature on internalism. For an early, but still useful, discussion of considerations that have been supplied in defence of internalism, see Frankena (1976 (1958)), pp. 49–73. For recent discussions, see e.g. Björnsson (1998),

(i) The consideration that most often is advanced in support of internalism is the intuition that it would be paradoxical if a person would *utter* a sentence to the effect that an action is right if she is not motivated to perform it.<sup>19</sup> For example, we would be perplexed were we to hear a person utter the sentence ‘It’s right to give money to the Red Cross’ and then come to believe that she is not at all motivated to give any money to the organisation. An important reason why we would find this paradoxical is, I think, that we have difficulties understanding *why* a person who is not so motivated would utter such a sentence. Unless she is accordingly motivated, we simply see no *point* in her uttering the sentence. Such an utterance, we might feel, would lack its appropriate function, and we would as a consequence be unsure how to understand it. As a result, in cases where a person utters a sentence to the effect that an action is right, but we have come to believe that she is not motivated to perform the action, we look for an explanation of why she nevertheless utters the sentence. We might for example come to believe that she does not, after all, judge that the action is right.

---

esp. chap. 3, and Cuneo (2002), pp. 465–486. For discussions of Hare’s argument, see e.g. Rønnow-Rasmussen (1993), chap. 1, and Björnsson (1998), pp. 55–60.

One traditional argument put forward in support of internalism should be mentioned, however. It runs, roughly, in the following way. It is claimed that if an action is right, there has to be a reason to perform it. It is further argued that according to internalism, moral judgements imply reason, but according to externalism they do not. Hence, it is concluded that internalism is correct and externalism incorrect. (For a recent expression of this argument, see Korsgaard (1996 (1989)), p. 43.) William K. Frankena was perhaps the first one to point out that the argument rests on an ambiguity as regards ‘reason’ between motivating reasons and normative reasons (Frankena (1976 (1958)), pp. 51–59). (For a recent version of this response, see Cuneo (2002), pp. 483–485.) True, internalists claim that if a person holds a moral judgement, she has a *motivating* reason to perform the action, whereas externalists deny this. But to hold this against externalism would be to beg the question. It might then be argued that the kind of reason the argument appeals to is not motivating reasons but normative reasons. However, neither internalism nor externalism claims that if a person judges that an action is right, it follows that she has a normative reason to perform the action. Both internalists and externalists can hold that the fact that an action is right implies normative reasons of various sorts, e.g. reasons of rationality. But this has nothing to do with them advocating internalism and externalism. Of course, if internalism is combined with the view that motivating reasons are normative reasons, it follows that if a person judges that an action is right, she has a normative reason to perform it. However, this is not something that follows from internalism itself. Moreover, the view that motivating reasons are normative reasons seems more controversial than internalism. It therefore seems implausible to presume this premise in an argument in favour of internalism.

<sup>19</sup> An utterance need not be verbal; it might be written or indicated in some other way.

This intuition seems to support internalism.<sup>20</sup> It is reasonable to argue that the explanation of why we feel that it would be paradoxical if a person would utter a sentence to the effect that an action is right unless she is motivated to perform it, is that a person's judgement which such a sentence is used to express entails that she is motivated to act in accordance with it.

However, it is important to note that this intuition does not imply that internalism is correct. One reason is of particular interest. Even if people do not *utter* a sentence to the effect that an action is right unless they are motivated to perform it, this is compatible with the possibility that they *judge*—but do not utter the corresponding sentence—that an action is right without being so motivated.<sup>21</sup> However, although the intuition under consideration does not imply that internalism is correct, it lends support to internalism since this view can explain it in the way just indicated.

(ii) Another consideration that is thought to support internalism is that we *expect* that when a person utters a sentence to the effect that an action is right, she is motivated to perform it.<sup>22</sup> This expectation is presumably vindicated by experience. It may be hypothesised that in most cases where we have experienced a person utter a sentence to this effect, she has either performed the action or at least shown some indications of being motivated to perform it. Accordingly, when this expectation is unfulfilled, we look for an explanation.

This consideration, more obviously than the previous one, does not imply that internalism is correct. But, again, it supports internalism in so far as internalism can explain it; the ultimate explanation for why we expect that a person's utterance to the effect that an action is right is accompanied by motivation to perform it might be that there is a connection between moral judgements and motivation of the sort maintained by internalists.

(iii) It may further be argued that we have the intuition that it would be paradoxical if a person would *judge* that an action is right if she is not motivated to perform it.<sup>23</sup> This intuition is the one of the three

---

<sup>20</sup> Philosophers attracted to internalism who take their point of departure in the philosophy of language seem especially influenced by this intuition; see e.g. Hare (1952), pp. 164–165, and Stevenson (1944), pp. 16–17. For other testimonies of this intuition, see e.g. Dancy (1993), p. 4; Lockie (1998), p. 16; Smith (1972), p. 87, and Smith (1994), p. 6.

<sup>21</sup> Cf. Frankena (1976 (1958)), pp. 66–67.

<sup>22</sup> See e.g. Smith (1972), p. 90.

<sup>23</sup> See e.g. Milo (1981), p. 375.

considerations that would give the strongest support to internalism. It is therefore interesting to note that it is seldom referred to in support of internalism.

One explanation as to why this intuition is rarely appealed to in support of internalism is presumably that it is weaker than the one mentioned in connection with the first consideration above (i). I think that we would find it truly paradoxical if a person would judge that an action is right, *utter* a sentence to this effect, but not be motivated to perform the action. However, we would find it less paradoxical if a person merely would judge that an action is right, *without* giving voice to her judgement, and not be motivated to perform the action. Consider the following two cases. Consider first a person who judges that it is right to give money to the Red Cross but who does not utter any sentence to this effect. She keeps, so to speak, her view to herself. However, try to imagine that she is not motivated to perform the action.<sup>24</sup> Consider next a person who judges that it is right to give money to the Red Cross and who also utters a sentence to this effect. We might for example imagine that she tells her friends that it is right to give money to the organisation, writes it in an e-mail or argues for this view in an article in a newspaper. However, try to imagine that she is not either motivated to give any money to the Red Cross. It seems that we would find the latter case more paradoxical than the former. Now, it might be argued that if internalism is correct, we should find the first case as paradoxical as the second. According to internalism, it is analytically necessary that if a person *judges* that an action is right, she is motivated to perform it. On this view, there does not seem to be any significant difference between the two cases that explains why we find the second case more paradoxical than the first one. The fact that we find the latter case more paradoxical may then be taken to suggest that our responses are not best explained by internalism, but by some other view which in certain respects has similar implications as this notion. This alternative view would then connect the motivation primarily to the utterance of the sentence in question and only secondarily and derivatively to the judgement. I will return to this issue in section 7.

---

<sup>24</sup> I use the expression 'try to imagine' here because it would be question begging against internalism to presume that we actually are able to imagine these situations.

This suggestion finds some support if we compare our reactions to the moral cases with our reactions to relevant non-moral cases. Consider first a person who judges that a certain individual is an uncle but who does not utter a sentence to this effect. Try to imagine that she holds that the individual in question is not a male. Consider next a person who also judges that a certain individual is an uncle and who does utter a sentence to this effect. Try to imagine that also this person holds that the individual in question is not a male. I think we find that the first case would be as paradoxical as the latter. And if we find the second case more paradoxical than the first one, the difference seems nevertheless lesser than between the moral cases. This suggests that *had* a person's judgement to the effect that an action is right with analytic necessity implied that she is motivated to perform the action, we would not have responded differently to the two moral cases either.

It should also be noted that even if the intuition under consideration—that it would be paradoxical if a person would judge that an action is right unless she is motivated to perform it—is accurate, it does not imply that internalism is correct. One reason is of particular importance. As we saw above, for internalism to be correct, the motivational state that motivates a person to do what she judges to be right has to be internal to her judgement, either in so far as her judgement consists in such a motivational state or in so far as it gives rise to such a state by itself. However, the intuition under consideration is compatible with that a person is motivated to do what she judges to be right, but that the motivational state in question is *external* to her judgement. But, again, even if this intuition does not imply that internalism is correct, it may gain support to internalism in so far as this view is able to explain it.

In connection to this observation, we should notice a further aspect of our intuitions that is relevant to the support the third consideration gives to internalism. We might find it paradoxical if a person would judge that an action is right without being motivated *at all* to perform it. However, I think we would find it less paradoxical if a person would judge that an action is right, *be* motivated to perform it, but be motivated merely by a motivational state that is *external* to her judgement. Consider the following two cases. Consider first a person who judges that it is right to help a certain person in need. However, try to imagine that she is not at all motivated to

provide the help. Consider next a person who also judges that it is right to help the person in question. Try to imagine, however, that she *is* motivated to provide the help, but that she is motivated by purely egoistic motives and not at all by any moral motives. That is, try to imagine that the motivation is external and not internal to her judgement. She may for example be motivated by a desire to promote her career or to avoid social ostracism.<sup>25</sup> It seems that we would find the first case more paradoxical than the second. Now, analogous to the previous cases, it might be argued that if internalism is correct, we should find the second case as paradoxical as the first one. Internalism says that if a person judges that an action is right, she is motivated to perform the action, where this motivation is *internal* to her judgement. According to internalism, there seems to be no significant difference between the two cases that explains why we find the first case more paradoxical than the second. The difference in our responses to the two cases can therefore be taken to indicate that our intuitions are not best explained by internalism, but rather by some other view the implications of which are in certain respects similar to those of internalism. According to this alternative view, motivation would be connected to the moral judgement irrespective of whether it is internal or external to it. I will return to this issue in section 7.

The intuition that would give the strongest support to internalism is that it would be paradoxical if a person would *judge* that an action is right if she is not motivated to perform the action by a motivational state that is *internal* to her judgement. It is therefore significant that advocates of internalism do not appeal to this intuition in defending their view.

If what I have argued in this section is correct, none of the considerations I have discussed implies that internalism is correct. As far as I understand, this means that the support these considerations give to internalism is weaker than often has been thought to be the case. They support internalism to the extent that internalism is able to explain them, but this is compatible with there being other explanations that are as good as, or even better than, the ones internalism offers. We have also seen that

---

<sup>25</sup> Here we have to presume that the person's 'purely egoistic' considerations are not part of her moral considerations. Consequently, we have to presume that her desire to do what promotes her career or to avoid social ostracism is not internal to her moral judgement. Note that we do not have to assume that egoistic considerations *are* not moral considerations. The only thing we have to assume is that they are not moral considerations according to *her* moral view.

there are some reasons to believe that there actually are superior explanations; in section 7, I will sketch the contours of such an account. Also when it comes to its explanatory power, internalism seems consequently weaker than it often has been assumed to be. However, it should immediately be recalled that there is an argument that may seem to provide a strong case for internalism: Smith's fetishist argument. This argument requires a separate treatment and discussion of it will be postponed to section 8.

### 5. *The Amoralist and Her Critics*

In the last section, I examined various considerations that are taken to support internalism. In this section and the next, I will examine whether there are reasons to believe that internalism is mistaken. The most obvious mode of doing so is to explore whether there are any counterexamples to it.

As is well known, externalists have argued that there is a kind of person who constitutes such a counterexample: *the amoralist*. An amoralist, as I will understand her, is simply any person who would falsify internalism by not being motivated in the way required by this view. Thus understood, the amoralist is a rather abstract figure whose character can be specified in different ways. The author whose discussion of the amoralist has become most influential is presumably David O. Brink. Brink, who argues against internalism by means of the amoralist, characterises such a person as one 'who is indifferent to what he concedes are moral considerations'.<sup>26</sup> In the next section, I will try to formulate versions of the amoralist which, in my view, make a stronger case against internalism than Brink's formulation might seem to allow, but for the time being we can bear Brink's characterisation in mind.

As can be suspected from that mentioned in the previous section, the amoralist is, at least at first appearance, likely to strike us as a quite peculiar character. Think of the amoralist as Brink sees her. We probably feel that it would be odd if a person would concede that a certain action is right but nonetheless not be motivated to perform it. Externalists can respond by admitting that the amoralist might not actually exist, and perhaps never have existed or will exist, but maintain that it nevertheless is possible to *conceive* of such a person. Advocates of this view can then maintain that since

---

<sup>26</sup> Brink (1989), p. 48.

internalism is claimed to be analytically necessary, the fact that we can conceive of the amoralist is sufficient to show that internalism is mistaken.<sup>27</sup>

Although the debate between internalists and externalists may be taken to indicate that people's intuitions about the amoralist diverge to some extent, I think it is difficult for internalists to outright deny that people take the amoralist to be conceivable.<sup>28</sup> That the amoralist appears to have this characteristic is something granted also by internalists.<sup>29</sup> Internalists do thus better to admit that the amoralist may *seem* conceivable, but argue that this impression is mistaken. In order to make this contention plausible, they suggest various explanations as to why this impression is mistaken.<sup>30</sup>

First, internalists may argue that the person who seems to be an amoralist does not really *judge* that the action in question is right, although it might appear so. Since she does not hold the judgement, internalists are not committed to claiming that she is motivated to perform the action. As this response stands, it is question begging since it merely denies something that is implied by the description of the amoralist. In order to be persuasive, it has to be supplemented with a reasonable explanation of why the person in question does not hold the judgement. Basically two such explanations have been offered.

According to R. M. Hare, we should understand the supposed amoralist as using the moral term in question in 'inverted commas'.<sup>31</sup> Applied to 'right', this means that, rather than judging that an action is right, she judges that *according to a moral view accepted by certain people*, the action is right. Thus understood, her judgement is not genuinely moral, but rather a judgement *about* how people who embrace a certain moral view would judge an action in terms of rightness. In contrast to genuine moral judgements, such a judgement does not entail that she is motivated to perform the action. This explanation might be supplemented by various stories as to why the person in question uses 'right' in this way.

According to another explanation of this kind, the supposed amoralist does not hold a judgement to the effect that an action is right because she is not linguistically competent with respect to 'right'. According to Michael

---

<sup>27</sup> Cf. Brink (1989), p. 46.

<sup>28</sup> See e.g. Smith (1994), p. 67.

<sup>29</sup> Cf. Smith (1994), p. 67.

<sup>30</sup> For an overview of different explanations internalists can propose, see Lockie (1998), p. 21.

<sup>31</sup> See e.g. Hare (1952), pp. 124–25, 164–167.



Smith, ‘the *very best* we can say about amoralists is that they try to make moral judgements but fail’, and the reason why they fail is that they lack ‘mastery of moral terms’.<sup>32</sup>

Second, internalists may argue that the person who seems to be an amoralist does not provide a counterexample to internalism because even if she does judge that an action is right, her motivation to perform it is cancelled because she is practically irrational. According to this view, the most plausible form of internalism incorporates a requirement of practical rationality in the way explained in section 3. When a person is not motivated to do what she judges to be right, this is accounted for by claiming that she suffers from some state that qualifies as practical irrationality.<sup>33</sup>

These explanations may seem persuasive. The reason is presumably our initial feeling that the amoralist is such a peculiar character. Each of the arguments gives an explanation as to why a person who seems peculiar *appears* possible to conceive of although this impression is mistaken. Our initial feeling that the amoralist is peculiar together with one or more of these explanations may thus have us conclude that our impression that such a person is conceivable is, on closer inspection, erroneous.

It might be argued, however, that this conclusion is premature. One reason is that it might be suspected that the explanations above succeed to dismiss the possibility of the amoralist only if she is described in certain particular ways. Take the explanation suggested by Smith, according to which the assumed amoralist is not linguistically competent in respect of ‘right’. It might be wondered whether it is not possible to describe an amoralist in such a way that we would consider her as competent in that respect. Also as regards the other internalist explanations it might be suspected that there are descriptions of amoralists in view of which we would come to similar conclusions. Consequently, it might be possible to argue that these explanations do not show that there is *no* amoralist who cannot be explained away in any of the manners suggested by internalists.

---

<sup>32</sup> Smith (1994), pp. 68, 70.

<sup>33</sup> Another strategy to explain away the assumed amoralist would be to argue that she *is* motivated to do what she judges to be right, but that her motivation is very weak. As far as I know, no one has appealed to this explanation. One problem with this response is that it seems question begging against externalism.

Let us now connect what I have said in this section with the discussion in the previous one. There it was observed that none of the considerations that typically are thought to support internalism implies it. Each consideration is consequently compatible with there being a person who constitutes a counterexample to internalism. In this section, we have found that the explanations internalists offer of why the amoralist erroneously appears conceivable may seem persuasive, but that it might be possible to provide descriptions of amoralists that could not be dismissed in any of these ways. These reflections suggest that there might be descriptions of amoralists such that they are both compatible with the internalist considerations and cannot be dismissed in the ways suggested by advocates of internalism. I will call amoralists who fulfil these conditions *advanced amoralists*. In the next section, I will indicate how amoralists of this sort can be understood. In the previous section, we made observations that indicate some of the characteristics advanced amoralists might have; I will accordingly return to these observations below.

## 6. *Advanced Amoralists*

In order to be able to characterise advanced amoralists, I will begin by making five observations regarding how amoralists may be described. Consideration of each observation will suggest how advanced amoralists can be understood.

(a) First I would like to call attention to an ambiguity that is common in descriptions of the amoralist. This ambiguity is illustrated by Brink's characterisation, according to which the amoralist is a person who is indifferent to what she *concedes* are moral considerations. According to one way of understanding the amoralist, she judges that an action is right, acknowledges this by uttering a sentence to this effect, but is not motivated to perform the action. According to another way of understanding her, she judges that an action is right, does *not* make any corresponding utterance, and is not motivated to perform the action. Formulations such as Brink's are open to both these interpretations. It is important to note that the latter formulation is sufficient for internalism to be mistaken. In section 4, we saw that the former case seems more paradoxical than the latter. It might therefore be suspected that one reason why the amoralist seems as such a peculiar person is that we think of her in the former way and not in the

latter. Moreover, understood in the latter way, the amoralist is compatible with the first internalist consideration.<sup>34</sup> This indicates that descriptions of advanced amoralists may adopt this alternative.

This observation is relevant to Hare's explanation mentioned above, according to which a supposed amoralist does not hold a genuine moral judgement but uses the moral term in question in inverted commas. According to this explanation, when the supposed amoralist uses 'right', she does not hold a genuine moral judgement, but rather a judgement about how people who embrace a certain moral view would judge an action in terms of rightness. In order for this explanation to be applicable, we have to imagine that she *uses* 'right', and hence that she *utters* a moral sentence involving this term. This explanation exploits the fact that when a person utters a sentence she may hold another judgement than the one the sentence ordinarily is used to express. This explanation is not applicable if we think of the amoralist in the latter way mentioned above. As the amoralist on that alternative does not utter a moral sentence, she cannot be dismissed by claiming that she uses the moral sentence to express another judgement than the sentence ordinarily is used to express.

(b) There is a further significant ambiguity in the formulation of the amoralist that is also illustrated by Brink's characterisation of this character. According to one way of understanding the amoralist, she judges that an action is right, but is not motivated to perform it by *any* motivational state whatsoever. This interpretation of the amoralist is predominant in the literature. According to another way of understanding her, she judges that an action is right, *is* motivated to perform it, but is motivated by a motivational state that is *external* to her judgement, not by any motivational state that is internal to it. Both these formulations are compatible with formulations such as Brink's. It is important to notice that the latter formulation is sufficient for internalism to be mistaken. As was noticed in section 4, we find the former case more paradoxical than the latter. It might therefore be suspected that one reason why we find the amoralist so peculiar is that we think of her in the former way and not in the latter. Moreover, understood in the latter way, the amoralist is compatible with the third

---

<sup>34</sup> This description of the amoralist is also compatible with the second internalist consideration.

internalist consideration. Hence, advanced amoralists may be understood in accordance with this alternative.

This observation is relevant to Smith's suggestion that the supposed amoralist does not hold a judgement to the effect that an action is right because she is not linguistically competent in respect of 'right'. According to the latter characterisation of the amoralist, she might be able to identify the actions to which 'right' applies and have her motivation guided in accordance with these identifications, although her motivation does not come from an internal motivational state but from an external motivational state. Note first that the fact that a person has the capacity to identify the objects to which a term applies in itself indicates that she is linguistically competent in respect of the term.<sup>35</sup> Hence, it is reasonable to hold that the amoralist, as characterised in the latter way above, is linguistically competent with respect to 'right'. However, she is not only able to do this; she also manages to have her motivation guided by what she takes 'right' to apply to. That is, she is motivated to do what she judges to be right. On the view considered in section 4—that we feel that there is a significant connection between judgements to the effect that actions are right and motivation—this should reinforce our view that she is linguistically competent as regards 'right'. Hence, if we understand the amoralist in the latter way, Smith's explanation seems less compelling. Of course, it might be objected that the person in question is not linguistically competent in respect of 'right' because the motivational state that accounts for her motivation is not internal to her judgement. But this contention is question begging since it presumes that being so motivated is part of the mentioned competence. Moreover, as we saw in section 4, internalists do not appeal to the intuition that motivation is internal to the pertinent judgements to support their view—presumably because we do not have this intuition.<sup>36</sup>

---

<sup>35</sup> Cf. Brink (1989), pp. 46–48; Smith (1994), pp. 66–67, and Svavarsdóttir (1999), pp. 191–192.

<sup>36</sup> The present point is also relevant for another argument put forward by Smith to the effect that a supposed amoralist does not hold a genuine moral judgement (Smith (1994), pp. 68–71). The argument consists in an analogy between colour judgements and moral judgements. According to Smith, 'the ability to have the appropriate visual experiences under suitable conditions is partially constitutive of possession of colour concepts and mastery of colour terms' (Smith (1994), p. 69). Hence, even if a colour blind person applies colour terms correctly, she does not hold colour judgements because she lacks the colour experiences that are constitutive of mastery of such terms. Similarly, according to Smith, even if a person applies 'right' to actions that have this property, she does not hold moral judgements because she lacks the motivation that is constitutive of mastery of

This observation also has implications for the way of dismissing the supposed amoralist that maintains that even if she does judge that an action is right, her lack of motivation is explained by the fact that she is practically irrational. This explanation is not applicable if we understand the amoralist in the latter way mentioned above. The reason is that she *is* motivated to do what she judges to be right, although not by a motivational state that is internal to her judgement, but one that is external to it. It might be objected that unless a person is motivated by a motivational state that is internal to her judgement, she is practically irrational. This objection raises difficult issues that cannot be dealt with here. However, it should be noted that this assumption does not find support in the observation which motivates that a requirement of practical rationality is incorporated into internalism. As we saw in section 2, the reason Smith adopts such a requirement is the observation that even if a person judges that an action is right, she might not be motivated to perform it because she suffers from, in Smith's words, 'weakness of will and other similar forms of practical unreason'. This observation might be taken to support the view that if a person is not motivated *at all* to do what she judges to be right, the explanation might be that she is practically irrational. This is also the implication Smith takes the observation to have. However, it does not support the view that if a person is not motivated to do what she judges to be right by a motivational state that is *internal* to her judgement, she is practically irrational. Neither is this something Smith maintains or provides arguments for.

(c) As mentioned in the last section, although the amoralist has been characterised in different ways by different authors, she appears usually as a rather abstract figure in the literature. One aspect of this abstractness is that we are not provided with any explanation as to *why* she lacks the appropriate moral motivation to do what she judges to be right. One reason why we find the amoralist peculiar may consequently be that we have not

---

this term. One difficulty with this analogy is that it seems doubtful whether colour experiences are constitutive of mastery of colour terms. However, the main difficulty is that according to the latter way of understanding the amoralist mentioned above, she *is* motivated to do what she judges to be right; it is just that she is not motivated by a state that is internal to her judgement, but by a state that is external to it. Hence, if we understand the amoralist in this way, she is analogous to the competent user of colour terms. This seems to mean that the analogy cannot be used to show that the supposed amoralist is not linguistically competent with respect to 'right' and hence that it cannot be used to show that she does not hold the relevant moral judgements. For other arguments against Smith's analogy, see Brink (1997), pp. 22–26, and Tenenbaum (2000), pp. 114–116.

been given such explanations. Hence, in describing advanced amoralists we may provide such explanations.

It might be argued that there are a number of reasonable explanations of this kind.<sup>37</sup> According to one kind of explanation, a person believes that doing what she judges to be right does not serve her self-interest and therefore loses her moral motivation. According to another kind of explanation, a person's indifference or negative attitude towards matters belonging to a certain area makes her lose her moral motivation in connection with these matters. According to a third kind of explanation, a person feels that morality requires a great deal from her and, overcome by these demands, she loses her moral motivation. Yet another kind of explanation applies when a person who finds herself in extraordinary circumstances, e.g. extreme fear, hunger or stress, loses her moral motivation. Some explanations might consist in combinations of two or more of these types of explanations.

(d) One aspect of the abstractness of the amoralist, as she ordinarily appears in the literature, is that she is assumed to lack the appropriate motivation as regards *all* the actions she judges to be right. This point is connected to the previous one. According to the most plausible explanations, she may not lack the motivation as regards all these actions, but only as regards those belonging to a certain kind.<sup>38</sup> For example, she might be morally motivated in relation to actions that concern her family but not in relation to actions that concern her work.

This observation is relevant to the internalist explanation that employs the idea that the supposed amoralist is not linguistically competent in respect of 'right'. The person we now are considering may be such that she, at least as regards most kinds of actions, is able to identify the actions to which 'right' applies and have her motivation guided in accordance with these identifications. It then seems implausible to maintain that she is not linguistically competent in respect of 'right'. Moreover, since she, at least as regards most kinds of actions, is motivated to do what she judges to be right, it is less plausible to claim that she is practically irrational.

(e) A further aspect of the abstractness of the amoralist, as she appears in the literature, is that she seems to lack the appropriate motivation as

---

<sup>37</sup> For proposals of such explanations, see Lillehammer (2002), pp. 18–21; Shafer-Landau (2003), pp. 148–151, and Svavardsóttir (1999), pp. 191–192.

<sup>38</sup> Cf. Brink (1997), pp. 24–25; Lillehammer (2002), p. 19, and Sadler (2003), p. 74.

regards the actions she judges to be right *all the time*. This observation is connected to the two previous ones. According to the most plausible explanations, the amoralist may lack the motivation only *occasionally*.<sup>39</sup> If she lacks the motivation only as regards a certain type of actions, she might lack it as regards *these* actions only occasionally.

This observation has the same implications for the internalist explanations as the previous observation. If the person in question lacks the motivation in question only occasionally, it is less plausible to claim that she is not linguistically competent in respect of ‘right’ or that she is practically irrational.

The five observations above, (a)–(e), suggest how advanced amoralists can be characterised. However, every characterisation of an advanced amoralist does not need to comprise all these observations, and different characterisations of advanced amoralists may put emphasis on different aspects of them. There might in other words be a number of descriptions of advanced amoralists. This holds particularly since, as we saw in relation to (c), there are a number of explanations as to why an amoralist is not accordingly motivated. Consider the following example.

Amanda has worked as a bureaucrat at the tax authority for many years. When she began, she found it quite interesting and was enthusiastic about her tasks. However, she has gradually become less interested in what she does. In fact, she finds her tasks so uninteresting that her only motive for continuing working there is that she is afraid that she would have difficulties finding another job. Amanda exemplifies a not too uncommon psychological phenomenon. If a person loses interest in a certain area, or has negative associations in relation to it, she might, at least temporarily, become unmotivated to perform actions connected to that area, unless she believes that performing these actions will affect her in egoistic terms. Now, her work involves moral considerations of various sorts. Particularly, it involves moral considerations in relation to the cases of suspected tax avoidance she handles. Here is an example of how Amanda’s moral motivation in relation to these cases is influenced by the situation she finds herself in. She holds that it is right for her to investigate into all the facts relevant to a particular case. However, due to the way she feels about her work, she is not motivated to do so by any other motive than an egoistic one. Thus, she is motivated to investigate into all the facts relevant to the case because she wants to keep her job, and she believes that if she does not handle the case properly, she might lose it. She is consequently motivated to do what she judges to be right, but only by a motivational state that is external to her moral judgement.<sup>40</sup> However, in relation to many other matters, Amanda is motivated to do what she judges right by non-egoistic motives. In fact, not even when it comes to handling cases of suspected tax avoidance she is always motivated by purely egoistic considerations. For example, she is sometimes—perhaps most of the time—motivated to

---

<sup>39</sup> Cf. Lillehammer (2002), p. 19, and Shafer-Landau (2003), p. 146.

<sup>40</sup> Here we have to assume that Amanda’s egoistic motives are not internal to her moral judgements.

investigate into all the facts relevant to a case because she judges that it is fair, and therefore right, to do so.

I think that it is plausible to assume that we find Amanda conceivable. Moreover, since she is described in accordance with the above observations, there is reason to believe that she is compatible with the considerations that are taken to support internalism and that she cannot be dismissed in the ways suggested by internalists. There are thus reasons to believe that Amanda provides support to the view that internalism is mistaken. I also think that the situation which she finds herself in is one that we might find ourselves in at times. Due to certain circumstances, we may sometimes be motivated to do what we judge to be right merely by non-moral considerations even if we otherwise are motivated by moral ones. If this is correct, Amanda would not only be conceivable; she would exemplify a kind of person who actually exists. As I suggested earlier, there are also reasons to think that there are descriptions of other advanced amoralists that would characterise them in such a way that we would find them conceivable. It should be pointed out that I do not claim that we do not find Amanda and other advanced amoralists peculiar. However, as I will argue in the next section, this is something externalism can account for.

### *7. Externalist Explanations of Internalist Considerations*

In this section, I will suggest that externalism can explain the considerations that are taken to support internalism. Moreover, I will suggest that these explanations are superior to those internalism offers. These explanations will also make it possible for externalists to account for why we may find amoralists peculiar. As suggested above, we may find amoralists strange because they seem difficult to reconcile with the considerations that are assumed to provide support to internalism. If externalists are able to explain these considerations, they are also in the position to explain why amoralists may seem peculiar. In proposing these explanations, I will draw on some ideas in pragmatics. To provide a satisfactory underpinning of the explanations would however require a substantial discussion of these ideas, something I am not able to provide here. As a consequence, my suggestions are incomplete and in need of further investigation.<sup>41</sup>

---

<sup>41</sup> At a very late stage in the work with this thesis, I found that David Copp and Stephen Finlay develop views that are highly relevant to the topic of the present section and in



Before I turn to the explanations inspired by pragmatics, I would like to call attention to two other kinds of explanations that externalists may put forward. The various explanations should however not be seen as competitors but rather as complementing each other. One kind of explanation appeals to the fact that the process of making us act morally is an essential part of our upbringing and, more generally, of shaping us into social beings. To facilitate this process, various psychological means are made use of; for example, doing what is judged to be right is associated with self-respect and other positive emotions, whereas doing what is wrong is associated with negative emotions such as guilt and shame. In case the process has been carried out successfully, we tend to feel that we have to do what we judge to be right and are consequently motivated to do so. According to another kind of explanation, moral judgements are connected to concerns that have a strong tendency to motivate us, such as various forms of well-being. Thus, one reason why we are motivated to do what we judge to be right is that actions that are right are connected to concerns of this type.

The latter kind of explanation is related to issues I will deal with in the next two sections. As we have seen, externalists account for a person's motivation to do what she judges to be right in terms of a motivational state that is external to her judgement. There are at least two questions that may be asked about such a motivational state. First, it might be asked what *kind* of content it has. On the assumption that it is a desire, it might be asked if it is desire to do what is right or a desire to perform actions that have certain other features. Second, it might be asked *what* the content of such a motivational state consist in. On the assumption that the motivational state is a desire, it might be asked what is involved in a desire to do what is right. Similarly, if it is a desire to perform actions with certain other features, it might be asked which these features are. I will bring up these issues in the next two sections. However, I will only discuss them in so far as they are directly relevant to the plausibility of externalism.

Let us now turn to the first internalist consideration noticed in section 3 and how externalism may explain it. According to this consideration, we

---

certain respects similar to the views I defend here; see Copp (2001), pp. 1–43, and Finlay (2004), pp. 205–223. Unfortunately, I have not been able to take into account their views.

would find it paradoxical if a person would *utter* a sentence to the effect that an action is right unless she is motivated to perform the action.

I think there might be various views in pragmatics that can help to explain this and the other two internalist considerations. This is especially so since my points are quite general and presumably not necessarily connected to a particular theory. One philosopher who suggests explanations of this kind is John Searle.<sup>42</sup> However, like some other commentators, I am attracted to the idea that such explanations can be formulated in terms of Paul Grice's notion of conversational implicature.<sup>43</sup> Grice argues that we as competent language users recognise certain rules so as to contribute to the kind of conversation in which we are engaged in a way that serves the generally accepted purpose of the conversation. The fact that we are understood to do so generates conversational implicatures. Thus, if a person utters a certain sentence, her utterance will—on the assumption that those involved in the communication are competent in the required way—be understood to conversationally implicate certain things due to the generally accepted purpose of the conversation. Take the following simple example. Imagine that two persons are putting together a bookcase. One purpose of their conversation is to facilitate their work. Suppose one of them is about to put shelves into the bookcase and that she says to the other person 'The shelves are over there'. Given the purpose of the conversation, it might be assumed that her utterance conversationally implicates that she wants the other person to give her the shelves.<sup>44</sup> If it turns out that she does not want the shelves, her utterance will seem strange. On Grice's theory, the

---

<sup>42</sup> Searle (1969), chap. 6, esp. pp. 139, 144–145, 147–149, 152, 154, and Searle (1979 (1975)), pp. 30–57, esp. pp. 32, 39–40. For an early suggestion of how externalist may explain internalist considerations along these lines, see Frankena (1976 (1958)), pp. 64–65, 67.

<sup>43</sup> The main text by Grice is Grice (1989a (1975)), pp. 22–40, but see also Grice (1989b (1975)), pp. 41–57. Some authors suggest that internalist considerations can be explained in terms of conversational implicatures; see e.g. Milo (1981), pp. 373–377, and Tenenbaum (2000), p. 125. For two significant discussions of how the relation between moral judgements and motivation can be explained in terms of pragmatics, see Copp (2001), pp. 1–43, and Finlay (2004), pp. 205–223. Finlay makes use of Grice's notion of conversational implicatures whereas Copp puts his views in terms of Frege's notion of 'colouring'. See also Barker (2000), pp. 268–279. Walter Sinnott-Armstrong argues that 'ought' sentences conversationally implicate 'can': Sinnott-Armstrong (1984), pp. 255–259.

<sup>44</sup> Of course, it might be necessary to add some information about the situation for this claim to be true.

explanation is that she then would have said something that is not consistent with a proper recognition of the purpose of the conversation.<sup>45</sup>

Grice emphasises that what is conversationally implicated is not a part of the meaning of a judgement, since it is not implied by the content of *what* the persons says, but rather ‘implied’—i.e. conversationally implicated—by her *saying* it: ‘the implicature is not carried by what is said, but only by the saying of what is said’.<sup>46</sup> Thus, that the person in the example above wants the other person to give her the shelves is not implied by her judgement, but conversationally implicated by her utterance. Conversational implicatures can be connected to a wide range of sentences. However, certain sentences and words may have conversational implicatures connected to them quite regularly. In such cases, it might be assumed that the connections between them can become conventionalised, at least to a certain extent.<sup>47</sup> As regards such sentences and words, it is easy to presume that implicatures are connected to the contents of what is said, although this is not the case.

Now, it seems reasonable to assume that one of our foremost concerns when we are involved in conversations about moral issues is to get each other to know how we are motivated in relation to certain types of actions.<sup>48</sup> There are presumably various reasons why we have this concern. One reason may issue from the fact that moral matters typically are relevant to our well-being or, more generally, to things we care much about. Thus, one reason for this concern may be that how we are motivated in relation to moral matters influences our actions, which in turn might affect the well-being of others, the well-being of people for whom they care, and, more generally, the things they care about. Another reason why we have this concern may issue from the fact that in conversations about moral matters,

---

<sup>45</sup> The rule which the person has offended is presumably the maxim ‘be relevant’; see Grice (1989a (1975)), p. 27. This is also the rule which seems brought into play in the moral cases below.

<sup>46</sup> Grice (1989a (1975)), p. 39. Cf. Searle (1979 (1975)), pp. 39–40, 42–43.

<sup>47</sup> See Searle (1979 (1975)), pp. 31, 40–42, 49–50. See also Grice (1989a (1975)), p. 37, and Grice (1989b (1975)), p. 43. However, this assumption raises the complicated issue what relation holds between Grice’s notion of *conversational* implicatures and his notion of *conventional* implicatures. As far as I understand, the distinction between them is not sharp. As I describe matters here, it would perhaps be more accurate to characterise the implicatures that are connected to moral utterances and sentences as conventional implicatures than as conversational implicatures. This would presumably also be more in line with Searle’s view. However, in this sketch, I will formulate my points in terms of conversational implicatures.

<sup>48</sup> See Smith (1972), p. 89–90. See also Milo (1981), p. 376–377.

we typically do such things as advise, commend, and prescribe people to perform or not to perform actions. That is, we use moral language to exert various kinds of influences on the actions of those who we communicate with. To accomplish these things and hence to exert the intended influences, it might be assumed that it is essential that we get others to understand whether or not we want the actions in question to be performed. Accordingly, it is essential that we get others to know that we are motivated in relation to these actions in certain ways.

It might be hypothesised that as a consequence of the mentioned concern, a generally recognised purpose of moral conversations is to let each other know whether or not we are motivated to perform certain actions.<sup>49</sup> On this assumption, when a person utters a sentence to the effect that an action is right, her utterance conversationally implicates that she is motivated to perform that action. Suppose it is accepted that the connection between sentences and conversational implicatures might become conventionalised to a certain extent. Given the importance of the mentioned purpose of moral conversations, it might be assumed that such a sentence quite generally is connected to the conversational implicature in question and that the connection between them has become accordingly conventionalised.

We can now provide an explanation of the first internalist considerations—that we would find it paradoxical if a person would utter a sentence to the effect that an action is right unless she is accordingly motivated. Suppose a person's utterance of a sentence to the effect that an action is right conversationally implicates that she is motivated to perform the action. If she is not motivated to perform the action, her utterance is not consistent with a generally accepted purpose of moral conversations. As a consequence, we find her utterance paradoxical. This also explains our feeling that there is no *point* in her utterance, that it does not fulfil its *function*, that we have difficulty understanding *why* she utters it. On the assumption that the conversational implicature is connected to the moral sentence in such a regular way that the connection has become conventionalised to a certain degree, it is even more readily explained that we find her utterance perplexing.

---

<sup>49</sup> Searle observes the importance of the 'point or purpose' of conversations for what sentences convey: Searle (1969), pp. 149, 154.

It is important to notice that this explanation does not mean that a person cannot *judge* that an action is right—without uttering the corresponding sentence—unless she is accordingly motivated. For, as we have seen, that a person is motivated to perform the action is not implied by what she says, the judgement the sentence expresses, but conversationally implicated by her *uttering* the sentence. Hence, this explanation is compatible with externalism.

A significant feature of conversational implicatures is that they are *cancellable* in the sense that a person can forestall a conversational implicature that her utterance would otherwise have.<sup>50</sup> To illustrate, return to the person in the example above who utters ‘The shelves are over there’. Suppose that just after having said this she says ‘But I don’t want them’. After hearing the second utterance, we would accept that she does not want the shelves in spite of the fact that this is what we would otherwise take the first utterance to conversationally implicate. However, we would be perplexed by her second utterance given the purpose of the conversation. Proper implications follow necessarily and cannot be cancelled in this way. Suppose a person says ‘That person is an uncle’ and then adds ‘But the person is not a male’. Her latter utterance does not mean that uncles are not males, since this is something that necessarily follows from something being an uncle. The sentences she has uttered are consequently contradictory. It is therefore interesting to notice that moral utterances seem to work more like the first kind of utterances than the latter. Suppose a person says ‘It’s right to give money to the Red Cross’ and then adds ‘But I haven’t the slightest inclination to give any money to that organisation’. I think we would accept that, after having uttered these sentences, she has succeeded in informing us that she judges that it is right to give money to the Red Cross but that she is not motivated to do so. In particular, I do not think we would consider the sentences she has uttered to be contradictory, as internalism would suggest. However, given that moral conversations have the purpose suggested above, we would find her second utterance perplexing.

Let us now turn to the second internalist consideration—that we *expect* that when a person utters a sentence to the effect that an action is right, she is motivated to perform it.

---

<sup>50</sup> Grice (1989a (1975)), p. 39. Cf. Searle (1979 (1975)), pp. 39–40.

Given that the appropriate conditions are fulfilled, people who are competent language users know that when they utter a sentence to the effect that an action is right, their utterance conversationally implicates that they are motivated to perform the action. As a result, there is a significant correlation between the voicing of such sentences and motivation to act in accordance with them which explains our expectation.

According to the third consideration that is thought to support internalism, we would find it paradoxical if a person would *judge* that an action is right unless she is motivated to perform it.

For reasons supplied in section 4, this consideration provides a weaker support to internalism than the first consideration; it is thus not in such an immediate need of explanation. However, I think there is an explanation on the lines above. (A similar explanation is suggested by Ronald D. Milo.<sup>51</sup>) We may start by considering a non-moral case. Suppose a person says 'Smith has not been to prison yet'. It might be assumed that this utterance conversationally implicates that Smith runs the risk of being put in prison.<sup>52</sup> However, it is possible for a person to hold the judgement that Smith has not been to prison yet without believing that Smith runs the risk of being put in prison. But in that case it would be strange to say that she believes that Smith has not been to prison yet, since this, given the conversational implicature, would be to suggest that she does believe that Smith runs the risk of being put in prison. A similar reasoning applies to the moral case. According to the suggestion above, an utterance to the effect that an action is right conversationally implicates that the person in question is motivated to perform the action. Given this implicature, it would be strange to describe her as holding the judgement that the action is right unless she is motivated to perform it, since this would be to suggest that she *is* motivated to perform the action. That is why we have the impression that the kind of situation referred to in the third internalist consideration would be paradoxical.

There is a related explanation of the third internalist consideration. (A similar explanation is proposed by M. B. E. Smith.<sup>53</sup>) On the assumption that a generally accepted purpose of engaging in moral conversations is to get each other to know what we are motivated to do, it is reasonable to

---

<sup>51</sup> Milo (1981), p. 377.

<sup>52</sup> Cf. Grice (1989a (1975)), p. 24.

<sup>53</sup> Smith (1972), pp. 89–90.

assume that we primarily are interested in people's moral judgements in so far as these judgements have implications for what they are motivated to do. It is therefore paradoxical to say that a person judges that an action is right unless she is motivated to perform it, since one then has said something that is irrelevant given the purpose of moral conversations.

In chapter 1, I mentioned that non-cognitivists may argue that moral sentences express beliefs, although secondarily. The reasoning above suggests that cognitivists who are externalists in a corresponding manner may argue that moral sentences express non-cognitive states, although secondarily. According to this account, if a person utters a sentence to the effect that an action is right, her utterance conversationally implicates that she is motivated to perform the action. In that way, what the person says expresses that she is in a certain motivational state. On the assumption that such states consist in non-cognitive states, it can be assumed that what she says expresses that she is in a certain motivational non-cognitive state. However, it does so only secondarily. Echoing the non-cognitivist account mentioned in chapter 1, the reason why this is so might perhaps be understood in the following way. Suppose a person asserts a sentence in which 'right' figures, e.g. 'It's right to give money to the Red Cross'. According to the present view, when 'right' is used in such a sentence, it is not necessarily used to express a certain non-cognitive attitude. More exactly, it is not the case that unless the person in question has such an attitude, her use of the term is incorrect. However, this claim should be qualified. It holds merely if we understand 'correct' in a narrow sense. As we have seen, that a person expresses a motivational state is a function of her *uttering* a sentence involving the term, where the utterance conversationally implicates that she is motivated to perform an action. On this view, the attitude in question is thus connected to the uttering of a sentence, not to the meaning or content of the term. As a consequence, the person's use of the term might be correct even if she lacks the attitude. However, for reasons provided above, her use of the term is inconsistent with a generally recognised purpose of moral conversations. In a broader sense of 'correct', her use of the term does therefore not qualify as correct.

Thus far I have argued that the three main considerations thought to support internalism can be explained in a way that is compatible with externalism. However, in section 4, we saw that our responses to two

significant cases suggest that externalism may be in a better position to explain our intuitions than internalism.

Let us first recall the observation that we would find it more paradoxical if a person would judge that an action is right *and* utter a sentence to this effect without being motivated to perform it, than if she merely would judge that an action is right without being motivated to perform it.

If internalism were correct, it would seem that we should find the second case as paradoxical as the first one. As we do not, internalism seems to have difficulty explaining our different responses to the two cases. Externalism can give an explanation on the lines above. According to that account, there is a significant connection between uttering sentences to the effect that actions are right and being motivated to perform these actions, since the utterances conversationally implicate that the persons in question are thus motivated. However, the connection between the corresponding judgements and motivation is weaker. As we have seen, to say that a person judges that an action is right might be taken to suggest that she is motivated to perform the action. But this is because the conversational implicature already is in place and because there is an underlying understanding of the purpose of moral conversations that generates this implicature. The judgement itself does not conversationally implicate anything of the kind; indeed, judgements do not conversationally implicate anything at all. Moreover, on this account, a person's judgement to the effect that an action is right does not, contrary to what internalism maintains, with analytic necessity imply that the person is motivated to act in accordance with it. Given this difference in strength between, on the one hand, uttering a sentence to the effect that an action is right and being motivated and, on the other hand, judging that an action is right and being motivated, it is quite natural that we respond differently to the two cases under consideration.

Let us next turn to the observation that we would find it more paradoxical if a person would judge that an action is right without being motivated *at all* to perform it, than if she would judge that an action is right and be motivated to perform it, but merely by a motivational state that is *external* to her judgement.

As was the case in relation to the last observation, internalism seems to have difficulty explaining why we respond differently to the two cases.



Externalism can give the following account. Above it was suggested that a primary purpose of engaging in moral conversations is to get each other to know what we are motivated to do. Given this purpose, it is fairly unimportant from what source the motivation comes; what is important is *that* a person is motivated to perform a certain action, not whether the state that motivates her is internal or external to her moral judgement. As a result, if a person utters a sentence to the effect that a certain action is right, her utterance conversationally implicates that she is motivated to perform the action, irrespective of whether the state that motivates her is internal or external to the corresponding judgement. Hence, it would be paradoxical if a person would utter such a sentence if she were not motivated at all to perform the action. But it would be less paradoxical if she would utter the sentence and be motivated by an external motivational state and not by an internal one. Now, according to the explanation proposed above, to the extent we find it misleading to say that a person judges that an action is right unless she is motivated to perform the action, this is a reflection of the conversational implicature and the purpose of moral conversations that generates it. As a consequence, we have a similar view of the relation between moral judgements and external and internal motivational states as we have of the relation between moral utterances and internal and external motivational states. We therefore find it more paradoxical to say that a person judges that an action is right if she is not motivated at all to perform it, than to say that she judges that an action is right if she is motivated merely by a motivational state that is external to her judgement.

#### **8. *Smith's Fetishist Argument***

Thus far I have argued that there are mainly two reasons to reject internalism and accept externalism. First, amoralists—particularly advanced amoralists—provide counterexamples to internalism. Second, externalism can provide explanations of the considerations that are taken to support internalism and these explanations are preferable to those internalism offers. The conclusion that externalism is superior to internalism is however preliminary. Michael Smith has presented a much discussed argument against externalism and for internalism known as ‘the fetishist argument’. Until I have made plausible that this argument is mistaken, I am not in the position to conclude that externalism is superior to internalism.

As we will see, Smith's argument raises an important question: what is the content of the motivational states that account for our moral motivation to do what we judge to be right? Smith argues that externalists cannot provide a satisfactory answer to this question whereas internalists can. In this section, I will discuss how externalist should answer this question in particular consideration of Smith's argument. In the next section, I will continue by arguing that realism provides externalism with a better account of the pertinent motivational states than reductionism.

According to Smith, considerations concerning the amoralist are not sufficient to decide between internalism and externalism. An independent argument is required, and the fetishist argument, he believes, is such an argument.<sup>54</sup> It is a 'striking fact about moral motivation', Smith claims, 'that a *change in motivation* follows reliably in the wake of a *change in moral judgement*, at least in the good and strong-willed person'.<sup>55</sup> Thus, if a good and strong-willed person changes her judgement about which action that is right, she will become motivated to perform the action she judges to be right after the change and lose her motivation to perform the action she judged to be right before the change.<sup>56</sup> Now, the question is how this *reliable connection* between changes in moral judgement and motivation is best explained.

On Smith's view, the kind of state that accounts for a person's motivation to do what she judges to be right consists in a desire.<sup>57</sup> He sees two ways to understand such a desire: either as a desire *de dicto* to do what

---

<sup>54</sup> Smith has presented somewhat different versions of the argument. Here I follow mainly Smith's original formulation of the argument (Smith (1994), pp. 71–76), since it, to my mind, provides the strongest account of it. I have also been helped by two latter formulations of the argument: Smith (1996), pp. 175–184, and Smith (1997), pp. 111–117. My understanding of the argument has also gained from some of the critical comments it has given rise to: Brink (1997), pp. 26–29; Cuneo (1999); pp. 359–380; Dreier (2000), pp. 619–638; Lillehammer (1997), pp. 187–195; Olson (2002), pp. 89–95; Shafer-Landau (1998), pp. 353–358, and Svavarsdóttir (1999), pp. 194–215. Smith offers the argument mainly as an argument against externalism and for his practicality requirement. However, the argument, if correct, supports also internalism as this view is understood here; cf. Smith (1994), p. 72.

<sup>55</sup> Smith (1994), p. 71.

<sup>56</sup> This holds only *ceteris paribus* because there might be motivational states that are not related to her moral judgements that are such that she still is motivated to perform the action in question after the change in moral judgement. In what follows, I will take the *ceteris paribus* clause for granted.

<sup>57</sup> The reference to desires here should not be taken to imply that internalists or externalists are committed to the Humean theory of motivation. We might consequently think of the relevant motivational states as beliefs or desires that are generated by beliefs. For the sake of simplicity, I will however stick to 'desire'.

the person judges to be right or as a desire *de re* to do what she judges to be right. The two types of desires differ as to their contents. If the person is motivated by a desire *de dicto* to do what she judges to be right, her desire has a content that involves the concept of rightness; the concept of rightness figures as a part of the intentional content of her desire. According to this alternative, she is motivated to perform actions with the particular aim to do what is right. If the person is motivated by a desire *de re* to do what she judges to be right, her desire does not have a content that involves the concept of rightness. In having such a desire, she is motivated to do what she judges to be right, but the concept of rightness is not part of the content of her desire. According to this alternative, she does not desire to perform the actions in question with the particular aim to do what is right. Put metaphorically, in the first case she desires to perform actions because they are right; in the second case she does not. A desire *de re* to do what is judged to be right might consist in a desire to perform actions with certain morally relevant features that right actions are thought to have, e.g. to promote the well-being of certain people.

Internalists explain the reliable connection in terms of the internal connection they think holds between a person's judgement to the effect that an action is right and motivation. Before the change, the good and strong-willed person's judgement to the effect that a certain action is right entails that she is motivated to perform that action. After the change, her judgement to the effect that a certain *other* action is right entails that she is motivated to perform *that* action. And as the first entailment does not hold after the change in judgement, she has lost her former motivation. Since holding a moral judgement of the indicated kind is sufficient for being accordingly motivated, there is no need, on this view, to appeal to anything but the judgements to explain the reliable connection. Smith believes that advocates of internalism are free to insist that good and strong-willed persons are motivated by a desire *de re*, not by a desire *de dicto*, to do what they judge to be right.

According to externalism, a person's judgement to the effect that an action is right does not entail that she is motivated to perform the action. Such judgements are thus not sufficient to explain the reliable connection between changes in moral judgements and motivation in the good and strong-willed person. Externalists therefore have to refer to a motivational

state that is external to such judgements which, in combination with these judgements, explains the reliable connection. On Smith's view, the only kind of motivational state that can fill this function is a desire *de dicto* to do what is judged to be right. But this explanation is implausible, Smith argues, because good and strong-willed persons cannot plausibly be considered to be motivated by such a desire:

Good people care non-derivatively about honesty, the weal and 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, where this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue.<sup>58</sup>

Smith concludes that since internalism is able to give an explanation of the reliable connection in terms of a desire *de re* to do what is judged to be right, whereas externalism is committed to an explanation in terms of a desire *de dicto*, internalism provides a better explanation of the reliable connection than externalism.

I think it can be argued that Smith's fetishist argument against externalism is mistaken. I will start by arguing that Smith does not appreciate that externalism can explain numerous instances of the reliable connection by appealing to a desire *de re*. I will then continue by arguing that in the cases where externalism may have to refer to a desire *de dicto*, this is not a problem for this view.

It is important to first make clear that externalists can offer an explanation of many instances of the reliable connection in terms of a desire *de re* to do what is judged to be right. This kind of explanation runs in terms of a change in beliefs about which actions that have certain features that right actions are thought to have together with a desire *de re*, where this is a desire to perform actions that have these features.

To see how this kind of explanation works, consider an example of the reliable connection. Suppose a good and strong-willed person initially judges that it is right not to give any money to the beggars in the town centre, but changes her view and comes to judge that it is right to give money to them, and that she changes motivation accordingly: before the change she was not motivated to give any money to the beggars, but after the change she is motivated to do so, and her former resistance has vanished.

---

<sup>58</sup> Smith (1994), p. 75.

Suppose now that she believes that actions that are right have certain, as she sees it, morally relevant features. In the present context, it is not important which these features are. However, there seem to be reasons to believe that features of that kind normally are fairly particular and concrete, but also that views of what they consist in might vary among people. We may assume that the person in question believes that actions which are right are such that they provide support to people in need, do not make people feel miserable, etc.<sup>59</sup> Suppose further that she has a desire to perform actions that have these features. Externalists can now give the following explanation of this instance of the reliable connection. Before the change in moral judgement, when she judged that it is right not to give money to the beggars, she did not believe that giving money to the beggars had these features. However, for some reason she comes to believe that giving money to the beggars actually has these features. (She may for example come to believe that they will not buy drugs for her money, something she thought before.) Accordingly, after the change in moral judgement, when she judges that giving money to the beggars is right, she believes that giving money to them has these features. Given her desire to perform actions with these features, she was not motivated to give any money to the beggars before the change in moral judgement, whereas she is motivated to do so after the change, and her former resistance has disappeared.

It is important to be clear about the nature of the kind of desire referred to in this kind of explanation. The features that the person in question assumes to be morally relevant consist in features that she believes that actions that are right have. The desire is consequently a desire to do what she judges to be right. However, this is a desire *de re*, not a desire *de dicto*. Although she believes that these actions are right, her desire to perform them does not involve the concept of rightness as a part of its content; she is not motivated to perform them because they are right. On this kind of explanation, there is thus no need to assume that good and strong-willed people are motivated by a desire *de dicto* to explain the reliable connection. Hence, just as internalists, externalists can refer to a desire *de re*

---

<sup>59</sup> On realism, it may seem natural to assume that these features consist in non-moral properties that, according to the person in question, make actions right (where 'make' is understood in accordance with the realist formula). But these features do presumably not have to consist in such right-making properties. They may for example consist in the properties that are involved in the reference-fixing characteristics associated with 'right'.

in the form of a desire to perform actions with certain morally relevant features to explain the reliable connection.

To my mind, this kind of externalist explanation of the reliable connection seems quite natural. Why, then, does Smith believe that externalists have to appeal to a desire *de dicto* to explain the reliable connection? The answer is not entirely clear from what Smith says. But as far as I understand, the reason is that he has in mind instances of the reliable connection that, unlike the kind of cases just considered, involve changes in beliefs about the morally relevant features of right actions.<sup>60</sup> In Smith's vocabulary, the person in question has changed her view of what features make actions right.

Consider again the example of the reliable connection mentioned above: the good and strong-willed person who changes judgement as to whether it is right to give money to the beggars and whose motivation changes accordingly. But assume now that she simultaneously changes her view of the morally relevant features of right actions. That is, assume that she simultaneously changes her view of which the morally relevant features are that actions that are right have. On the kind of externalist account proposed above, the reliable connection is explained in terms of a change in belief about which actions that have the morally relevant features together with a desire to perform actions that have these features. These features remain the same before and after the change; the person in question has merely changed her view of *which actions* that have them. In the case we now are considering, the person has however changed her view of which these features *are*, not, or not merely, her view of which actions that have them.<sup>61</sup> The externalist account described above thus refers to a desire to perform actions with features that she after the change no longer believes that right actions have. It can consequently not figure in an explanation of why she after the change in moral judgement, when she judges that it is right to give money to the beggars, is motivated to act in accordance with

---

<sup>60</sup> Smith discusses an example where a person first decides to vote for the libertarian party and after a change in his 'most fundamental values' decides to vote for the social democrats (Smith (1994), pp. 71–72). He also considers a case where a person starts off being a strict utilitarian, but then becomes an advocate of a normative view that accepts special concerns for family members (Smith (1996), pp. 180–181). See also Smith (1997), pp. 113–115.

<sup>61</sup> On realism, it is plausible to assume that the person in question has changed her view of which the non-moral properties are that make actions right (where 'make' is understood in the accordance with the realist formula).

her judgement. It might therefore be concluded that externalism cannot provide explanations of instances of the reliable connection like this one in terms of a desire *de re*.

Externalists can provide explanations of these instances of the reliable connection in terms of a change in moral judgement together with a desire *de dicto* to do whatever is judged to be right. Return to the example above. Before the change in moral judgement, when the person judges that it is right not to give money to the beggars, she is, due to her desire *de dicto* to do what she judges to be right, not motivated to give any money to the beggars. After the change, when she judges that it is right to give money to the beggars, she is, due to her desire *de dicto*, motivated to do so.

If this reasoning is correct, externalists have to maintain that good and strong-willed persons are motivated by a desire *de dicto* to explain the instances of the reliable connection where these persons alter beliefs about the morally relevant features of right actions. It is important to see that this conclusion is weaker than the one Smith believes follows from the argument. As he presents the argument, it is supposed to show that externalists are committed to explaining *all* instances of the reliable connection in good and strong-willed persons in terms of such a desire. But, as we have seen, externalists can explain the instances of the reliable connection where good and strong-willed persons do not alter beliefs in the mentioned way in terms of a desire *de re*.

There is a further consideration which indicates that Smith's argument is weaker than it appears to be on first appearance. This issue concerns what is involved in being a 'good and strong-willed person'. Of course, what is involved in being such a person is a complicated issue; here I will only consider two simplified conceptions. According to a *substantial* conception of a good and strong-willed person, she is motivated to do what *in fact* is right, not merely what she *judges* to be right. This is presumably the commonsense understanding of such a person.<sup>62</sup> According to a *non-substantial* conception of a good and strong-willed person, she is motivated to do what she judges to be right, but not necessarily what in fact is right. Now, if we assume the substantial conception of what it means to be good and strong-willed, the reasoning directed against externalism above is mistaken. Since the good and strong-willed person on this conception is

---

<sup>62</sup> Cf. Copp (1997), p. 51.

motivated to do what in fact is right, there can be no question of her *changing* motivation in the relevant respect. In that case there is no reliable connection that is in need of explanation, and the argument above does not get off the ground.<sup>63</sup> Accordingly, in a subsequent comment on his argument, Smith makes clear that it is not the substantial conception he has in mind, but the non-substantial.<sup>64</sup>

There are at least two difficulties with Smith's non-substantial conception of a good and strong-willed person in the present context. One is that it does not accord with our notion of good people. Suppose, for instance, that a convinced racist believes that right actions are such that they contribute to the extermination of a certain race and that she is motivated to perform these actions. According to the non-substantial conception, she would qualify as good, but this result seems counterintuitive. However, the main difficulty is that if we understand a good and strong-willed person in the non-substantial way, it is not at all clear why it would be any problem for externalists to claim that such a person is motivated by a desire *de dicto* to do what she judges to be right. According to the substantial conception of the good and strong-willed person, it may perhaps seem problematic to assume that she is motivated by such a desire because it might be argued that

---

<sup>63</sup> This argument may need to be qualified. It might be the case that a substantially good and strong-willed person is not motivated to perform actions that are right. To put the point in terms I have utilised above, the reason is that she may be mistaken about which actions have the morally relevant features of right actions. That is, she may be mistaken about which actions have the morally relevant features that actions that are right have. (Cf. Stratton-Lake (2000), p. 16–17.) But something that she presumably is not wrong about is which these features *are*. For example, if helping people in need is such a feature, she recognises this and is motivated to perform actions which she believes have this feature. However, she might be mistaken about which actions have this feature and as a consequence be motivated to perform actions that are not right. Which morally relevant features actions that are right have is stated in the correct moral account of such actions. Smith claims accordingly that substantially good and strong-willed persons have 'the motivations that the one true morality tells them that they should have' (Smith (1996), p. 177). It is primarily in this sense that a good and strong-willed is motivated to do what in fact is right. As we have seen, the substantially good and strong-willed person recognises the mentioned features and changes accordingly not her view of which these are. Neither does she change her motivation in relation to these features. Hence, she does not change moral judgement or motivation in a way that externalists need to explain in terms of a desire *de dicto* to do what is judged to be right. Consequently, this qualification does not affect my point against Smith's argument.

<sup>64</sup> Smith (1996), p. 176–177. In order avoid misinterpretations of his argument, Smith has abandoned the term 'good and strong-willed person' and prefers instead the term 'moralist'.



persons with that character are not motivated by a desire that has the content of a desire *de dicto*. (However, I am not sure this is correct either.) This is also the idea suggested by much of what Smith says about good and strong-willed persons. Consider for example the quotation above where he claims that such a person is motivated by '*honesty, the well and woe of their children and friends*', etc., and not by a desire *de dicto*, because the latter would be a '*fetish or moral vice*'.<sup>65</sup> When Smith appeals to the content of a good and strong-willed person's desire in this way, he clearly has a substantially good and strong-willed person in mind. However, on the non-substantial conception of a good and strong-willed person, it is not at all that obvious why it would be so problematic to claim that she is motivated by a desire *de dicto*, as there does not seem to be anything in such a person's character that rules out that she is motivated by a desire with such a content.

If what I have argued so far is correct, the fetishist argument is weaker than it seems at first appearance. Externalists are not committed to explaining the reliable connection in substantially good and strong-willed persons by claiming that they are motivated by a desire *de dicto* to do what they judge to be right. The only instances of the reliable connection which externalists might need to explain in terms of such a desire are those that involve changes in beliefs about which the morally relevant features are in persons who are non-substantially good and strong-willed.<sup>66</sup> Would this be a problem for externalists?

---

<sup>65</sup> Smith (1994), p. 75. Italics added.

<sup>66</sup> It should be observed that externalists *are* able to provide explanations of at least some instances of the reliable connection which involve changes in beliefs about the morally relevant features of right actions in terms of a desire *de re* to do what is judged to be right. Externalists can do so on the assumption that the person in question has a pre-existing desire to perform actions with certain features and that she after the change comes to believe that these features are had by actions that are right. Let us consider an illustration of how this kind of explanation might work. Recall again the example above: the person who changes her judgement as to whether giving money to the beggars is right and who modifies her motivation accordingly, while she simultaneously alters her view as regards the morally relevant features of right actions. Before the change she believes that actions that are right have certain, as she sees it, morally relevant features. Let us call these features F1. To illustrate, we may assume that one member of F1 is that of contributing to people's ability of taking care of themselves. She has a desire to perform actions with F1. She does not believe that giving money to the beggars has F1 and is accordingly not motivated to give any money to them. We now add the assumption that the person in question has a desire to perform actions that have certain other features, F2. One feature of F2, we might assume, is that of contributing to equality among people. Now, the following happens. As she contemplates her moral view, she comes to doubt whether F1 really are the features right actions have. A result of this process is that she becomes convinced that F2, not F1, are the morally relevant features these actions have. As we saw, she has the desire to perform actions with F2.

I think the answer to this question is ‘no’. As far as I understand, we are quite commonly motivated by a desire *de dicto* to do what we judge to be right. Rather than defending the view that it is *not* problematic to refer to a desire *de dicto* to explain the indicated instances of the reliable connection, I will therefore suggest that it is difficult to see that there are any good reasons to believe that it *is* problematic; in particular, I will argue that it is difficult to see that Smith has offered any such reasons.

One important reason why it may be thought to be problematic to claim that we are motivated by a desire *de dicto* to do what we judge to be right is, as I have already suggested, due to a failure to distinguish between substantially and non-substantially good and strong-willed persons. When Smith starts off arguing against explaining the reliable connection in terms of a desire *de dicto*, he has a substantially good and strong-willed person in mind. He then continues his reasoning on the presumption that the same argument holds for a non-substantially good and strong-willed person. However, I fail to see that he has provided any reason for this view.

Another reason why it might be thought problematic is the view that a person who is motivated by a desire *de dicto* to do what she judges to be right has to be consciously aware that she is motivated by such a desire. It might be argued that this would be awkward because we do not experience ourselves as being motivated by a desire to do what is right, where this is understood as a desire *de dicto*. Smith indicates that he believes that externalists are committed to this conception when he says that, according to this view, a good and strong-willed person is motivated by a ‘self-consciously moral motive’.<sup>67</sup> However, it is difficult to see why externalists should be thus committed. Externalists may—and presumably should—claim that we are not consciously aware that we are motivated by a desire *de*

---

The change in her view of the features of right actions, has her reflecting over which actions that have F2. When she considers whether giving money to the beggars has F2, she realises that it actually has these features. She becomes accordingly motivated to give money to the beggars. This explanation of the reliable connection does not appeal to any desire *de dicto* to do what she judges to be right. Before the change, she was motivated by a desire *de re* in the form of a desire to perform actions that have features F1, and after the change she is motivated by a desire *de re* in the form of a desire to perform actions that have features F2. However, this kind of explanation rests on a crucial presumption: that the person in question has a pre-existing desire to perform actions with certain features and that she after the change comes to believe that these features are had by actions that are right. We can presumably not make this presumption as regards all relevant instances of the reliable connection.

<sup>67</sup> Smith (1994), p. 74.

*dicto*. They may in other words claim that such a desire stays in the background of our moral deliberations.<sup>68</sup>

There is a further reason why it may be thought to be problematic to claim that we are motivated by a desire *de dicto* to perform actions that we judge to be right. Some of the things Smith says suggest that he believes that when we are motivated by such a desire, we are motivated to perform these actions without any consideration of their properties other than that they are right. He writes in other words as if we on this alternative would be motivated to perform these actions irrespective of which other properties we believe they have.<sup>69</sup>

However, there seems to be no reason to assume that this is the case. On any reasonable view of rightness, it is the case that if a person judges that an action is right, she believes that the action has this property in virtue of having certain non-moral properties. This means that when she is motivated by a desire *de dicto* to do what she judges to be right, she is not motivated to perform these actions irrespective of which other features she believes they have. One way to see this is the following. Suppose a person is motivated to perform an action by a desire *de dicto* to do what she judges to be right. She then believes that the action is right in virtue of having certain non-moral properties. If she comes to believe that the action does not have the non-moral properties she first thought it had, she might come to believe that the action is not right after all. Given her desire *de dicto*, her motivation to perform the action will then vanish.<sup>70</sup>

If these remarks are correct, there is reason to believe that there is no problem for externalists to refer to a desire *de dicto* to do what is judged to be right in explaining the instances of the reliable connection mentioned above: those that involve changes in beliefs as regards which the morally relevant features are in persons who are non-substantially good and strong-willed. And, as we have seen, these are the only cases where externalists might need to refer to such a desire. Together with the arguments presented earlier, this provides reasons to reject Smith's fetishist argument.

---

<sup>68</sup> Cf. Shafer-Landau (2003), pp. 158–159, and Svavarsdóttir (1999), p. 202. For the distinction between background and foreground desires, see Pettit and Smith (1990), pp. 565–592.

<sup>69</sup> See e.g. Smith (1997), pp. 114–115.

<sup>70</sup> Cf. Zangwill (2003), pp. 148–149.

One important question I have left uncommented is whether there would be any difficulty for externalists in maintaining that also *substantially* good and strong-willed persons are motivated by a desire *de dicto* to do what they judge to be right. Like a number of other commentators on Smith's argument, I have difficulty seeing anything problematic about this.<sup>71</sup> Moreover, I am inclined to think that considerations such as those just offered can explain why it is mistakenly thought to be problematic. If this is correct, there is no reason to believe that the only instances of the reliable connection that are explainable in terms of a desire *de dicto* are those mentioned above. On the contrary, many, perhaps the majority, of the instances of the reliable connection might be explainable in this way. However, as Smith's discussion bears witness, some people seem to have the very strong intuition that substantially good and strong-willed persons are not motivated by such a desire. Hence, it might be difficult to argue against this view in a convincing way. And, as we have seen, it is not necessary to do so in order to cast doubt on Smith's argument.

#### 9. *Reductionism, Realism and Moral Motivation*

According to externalism, a person's motivation to do what she judges to be right consists in a motivational state that is external to her judgement. In the last section, we saw that on the assumption that motivational states are desires, externalism may suggest that such a motivational state consists in a desire with either of two types of contents. Either it is a desire *de re* to do what is judged to be right. That is, it is a desire to perform actions that have certain features which are assumed to be morally relevant, where these features are thought to be had by actions that are right. Or it is a desire *de dicto* to do what is judged to be right. That is, it is, briefly put, a desire to perform actions with the particular aim to do what is right.

I do not think that meta-ethical theories have anything significant to say about desires *de re* to do what is judged to be right. As I indicated in the last section, the features involved in such desires are presumably rather concrete and particular. Moreover, people might differ to a great extent as to which these features are. There seems to be little reason to believe that a

---

<sup>71</sup> Cf. Copp (1997), pp. 49–50; Sadler (2003), pp. 69–71; Svavarsdóttir (1999), pp. 202–203, and Zangwill (2003), pp. 146–148.

meta-ethical theory could capture this variety, and I see no particular reason why it should.

Concerning the desires *de dicto*, things are different. Since meta-ethical theories are, among other things, about the nature of moral properties, we can require that such a theory is able to account for this kind of desire. This means that one relevant measure of the plausibility of a meta-ethical theory is whether it succeeds in doing so. Moreover, for externalism it is especially important that there is a meta-ethical theory that can meet this requirement, since it explains certain instances of the reliable connection in terms of a desire *de dicto*. If externalism cannot be combined with a meta-ethical theory of this kind, there is in other words reason to believe that externalism cannot explain moral motivation in a satisfactory way.

Let us consider how analytic reductionism, synthetic reductionism and realism, respectively, characterise a desire *de dicto* to what is judged to be right.

#### *Analytic reductionism*

Analytic reductionism has a significant consequence for the content of a desire *de dicto* to do what is judged to be right. According to this view, the meaning of 'right' is constituted by a certain non-moral property. As we have seen earlier, a person who is linguistically competent as regards the meaning of 'right' knows what constitutes the meaning of the term. This suggests that if she is motivated by a desire *de dicto* to do what she judges to be right, she desires to perform actions that have the mentioned property. Thus, the property will be part of the content of her desire *de dicto*.<sup>72</sup>

Let us next note that on any version of analytic reductionism that appears at least remotely plausible, the non-moral property that constitutes the meaning of 'right' presumably will be quite abstract. This is a consequence of various requirements placed on a correct reductive account of the meaning of the term. According to one of these requirements, this property has to be recognised by anyone who is linguistically competent in respect of the meaning of 'right' and hence by people who embrace widely different substantive normative conceptions. Consequently, according to an

---

<sup>72</sup> Of course, she does not need be consciously aware that she is motivated by a desire with this content.

influential family of views, the meaning of ‘right’—and other ‘thin’ moral terms—should be understood in terms of what a highly idealised creature, e.g. an ideal observer, would approve of.<sup>73</sup>

However, I suspect that the non-moral property that constitutes the meaning of ‘right’ on any plausible account is so abstract that it is, at least generally, implausible to ascribe to us a desire to perform actions that have such a property. For example, it does not seem to be an accurate description of our moral motivation to claim that we are motivated by a desire to perform actions that a highly idealised creature would approve of. Admittedly, we may at times be motivated by such a desire. But it seems implausible to suppose that we are motivated by such a desire whenever we are motivated by a desire *de dicto* to do what we judge to be right.

It might also be wondered *why*, from our own perspective, we would be motivated by a desire to perform actions that have such an abstract property. This question arises particularly as it might be very difficult for us to get to know which actions that have this property. For example, suppose that the meaning of ‘right’ is understood in terms of what a highly idealised creature, such as an ideal observer, would approve of. Such a creature presumably has characteristics that no existing human being has; for example, she has a unique epistemic position, involving all relevant true beliefs, etc.<sup>74</sup> As a consequence, it will be quite hard, perhaps impossible, for us to know which actions such a highly idealised creature would approve of. But then it seems difficult to understand why we would be motivated by a desire to perform these actions. After all, as far as we know, the actions such a creature would approve of might be ones that we are not at all interested in seeing carried out. Moreover, given that it is quite difficult to get to know which actions that have the non-moral property that constitutes the meaning of ‘right’, it is quite difficult to know which actions to perform in order to do what is right. It would then be even harder to understand why we would be motivated to perform actions that have this property.

Gilbert Harman has pressed a related objection against the view that people who are motivated to do what is right are motivated by a desire to perform actions that an ideal observed would approve of. Harman argues that such a desire would ‘too “outer directed” to count as a moral

---

<sup>73</sup> For an elaborated version of this view, see Firth (1952), pp. 317–345.

<sup>74</sup> Cf. Firth (1952), pp. 333–335.

motive'.<sup>75</sup> If I understand him correctly, he thinks that a person who is motivated by such a desire would be too much concerned with what an ideal observer would approve of and too little concerned with the qualities of her own actions. Harman concludes: 'Such a desire is precisely not a desire to do something simply because it is right.'<sup>76</sup>

However, it should be admitted that whether these arguments are plausible depends on what the meaning of 'right', according to analytic reductionists, consists in. There might be an account of the meaning of 'right' that evades them. Hence, they are by no means conclusive.

### *Synthetic reductionism*

Synthetic reductionism does not have the same consequence as analytic reductionism regarding the content of a desire *de dicto* to do what is judged to be right. According to this view, the reference of 'right' consists in a certain non-moral property. However, it is not part of being linguistically competent with respect to 'right' to know what the reference of the term consists in. This means that when a person has a desire *de dicto* to do what she judges to be right, it does not follow that she desires to perform actions that have a certain non-moral property, where this property constitutes the actual reference of 'right'. Hence, this property need not be part of the content of her desire *de dicto*.

This view does not have the difficulty with analytic reductionism mentioned above. Since it does not imply that a person who is motivated by a desire *de dicto* is motivated by a desire whose content involves the non-moral property that 'right' refers to, there is no reason to believe that the content of her desire involves an unduly abstract property.

However, synthetic reductionism seems to have a problem that is relevant in the present context. As just mentioned, on this view it is not part of a person's linguistic competence in respect of 'right' to know what non-moral property constitutes the reference of the term. She refers to this non-moral property even if she is not aware of it and even if she believes that rightness consists in something else entirely. This has the consequence that there may be an awkward discrepancy between the content of the moral sentences a person utters and her moral motivation.

---

<sup>75</sup> Harman (1986b), p. 1.

<sup>76</sup> Harman (1986b), p. 4. Cf. Johnston (1989), pp. 157–158.

To see this, consider the following example. Assume that the non-moral property that 'right' refers to according to synthetic reductionism is that of maximising happiness. When a person utters a moral sentence such as 'It's right to give money to the Red Cross', 'right' thus refers to maximising happiness. However, imagine a person whose view of rightness does not accord with what 'right' actually refers to and who believes that rightness consists in something else than maximising happiness. We may for example think of a convinced defender of a certain normative conception, such as the dedicated deontologist mentioned in chapter 5. Assume further that she is motivated to perform the action by a desire *de dicto* to do what she judges to be right. Now, what she desires to do is what she judges to be right, where this feature consists in what she *believes* rightness consists in. What she desires to do is what is right according to *her* view of rightness. As a consequence, in the example at issue, the feature which is involved in the person's desire *de dicto* does not correspond to what 'right' actually refers to.

As a result, when the person in question utters 'It's right to give money to the Red Cross', 'right' in this sentence refers to another feature than the feature that is involved in the content of her desire *de dicto* to do what she judges to be right. This consequence seems problematic. When a person utters a sentence of this type, and she is motivated to perform the action, we take 'right' to indicate a feature of the action which is relevant to her being motivated to perform it. More precisely, we take the term to indicate a feature that helps to explain the person's motivation. This seems especially to be the case when we take her to be motivated by a desire *de dicto*. However, in the mentioned kind of cases, the property referred to in the sentence may not be relevant to her motivation to perform the action. Hence, it does not have the mentioned explanatory function. There is one kind of cases where this consequence seems particularly awkward. I have in mind the cases where the sentence in question is true, where the person consequently is correct in her belief that the action is right, and where she is motivated by a desire *de dicto* to do what she judges to be right. In such cases we seem to assume that in referring to 'right', she has succeeded in identifying a feature of the action, rightness, where this feature is such that it helps to explain her motivation to perform the action. However, as far as I see, this conception cannot be maintained on synthetic reductionism.



### *Realism*

We may start by observing that realism does not have the implication of analytic reductionism concerning the content of a desire *de dicto* to do what is judged to be right. Realism denies that the meaning of ‘right’ is constituted by a particular non-moral property. Hence, the content of a desire *de dicto* does not involve such a property. According to the realist formula, the relevant aspect of the meaning of ‘right’ can instead be captured by saying that an action to which the term applies has a set of non-moral properties—some set of non-moral properties—such that it makes actions right. (For the sake of simplicity, I put my points here in terms of the simplest formulation of the realist formula.) Hence, in virtue of being linguistically competent in respect of the meaning of ‘right’, a person who is motivated by a desire *de dicto* desires to perform actions that have this property. This property will consequently be part of the content of her desire *de dicto*.

Since the content of a desire *de dicto* according to realism does not involve any particular non-moral property that constitutes the meaning of ‘right’, there is no reason to believe that this view characterises the content of the desire as too abstract, at least not in the way that provides a difficulty for analytic reductionism.

As we saw earlier, synthetic reductionism avoids this difficulty. However, realism avoids also the difficulty with synthetic reductionism mentioned above. To see this, consider an example that corresponds to the one discussed earlier. Suppose a person utters a moral sentence such as ‘It’s right to give money to the Red Cross’. According to realism, ‘right’ in this sentence does not refer to a particular non-moral property. Rather, it refers, in accordance with the realist formula, to the irreducible property rightness where an action with this property has a set of non-moral properties such that it makes actions right. Let us assume that the set of non-moral properties which makes the action in question right is that of maximising happiness. However, imagine a person who believes that what makes actions right is something else than actually makes them right, i.e. something else than maximising happiness. Her conception of what makes actions right is in other words erroneous. Moreover, assume that this person is motivated by a desire *de dicto* to do what she judges to be right. Now, on the realist formula, the relevant aspect of the meaning of ‘right’ can be

captured by saying that an action to which the term applies has a set of non-moral properties such that it makes actions right. This means, as we saw above, that in virtue of being linguistically competent with respect to the meaning of ‘right’, a person who is motivated by a desire *de dicto* to do what she judges to be right is motivated by a desire to perform actions that have precisely this feature. Consequently, the feature that is involved in her *desire de dicto* to do what she judges to be right corresponds to what ‘right’ refers to. And this holds irrespective of her being mistaken about what makes actions right.

On this view, the discrepancy between the content of moral sentences and moral motivation that troubles synthetic reductionism does not emerge. That is, on realism there is no discrepancy between, on the one hand, what ‘right’ refers to in the moral sentence the person utters, and, on the other, the feature which is involved in the content of her desire *de dicto* to do what she judges to be right. Even though she is mistaken about which non-moral properties that make actions right, ‘right’ in the sentence she utters refers to the same feature that is contained in her desire *de dicto*, i.e. rightness as this property is understood according to the realist formula. Consequently, realism does not have the awkward consequence of synthetic reductionism.

## 7. Concluding Remarks

In this chapter, I have considered an important argument against cognitivism and in favour of non-cognitivism: the internalist argument. I have maintained that this argument can be rejected by denying one of its premises, internalism, and by accepting externalism. First, I argued that there are counterexamples to internalism in the form of amoralists—particularly in the form of what I have called advanced amoralists. Second, I argued that there is reason to believe that externalism can explain the considerations that are thought to provide support to internalism and that these explanations are preferable to those offered by internalism. (It is worth pointing out, however, that the first line of reasoning, if successful, is sufficient to show that internalism is mistaken.) In the latter line of reasoning, I appealed to some ideas in pragmatics. However, this account is incomplete and in need of further investigation. Moreover, I maintained that Michel Smith’s fetishist argument in favour of internalism and against

externalist fails. Smith's argument raises an important question that externalists have to address, namely how moral motivation should be understood on this view. I finished the chapter with arguing that realism provides a better account than reductionism of certain pertinent instances of moral motivation. As a consequence, realism, unlike, reductionism, is able to contribute to the defence of cognitivism against the internalist argument.

## Chapter 8

### Moral Dependence

#### 1. *Introduction*

The distinguishing characteristic of realism, as this view is understood here, is that moral properties constitute a separate category of properties and that they consequently are irreducible to non-moral properties. In chapter 4, I pointed out that notwithstanding their stress on the irreducibility of moral properties, realists have to be able to account for our notion that objects have moral properties in virtue of having non-moral properties. I hypothesised that realists can do so by adopting the realist formula. This formula, I maintained, gains support from our considerations in relation to the questions posed in chapters 2 and 3. In the three subsequent chapters, I then argued that realism, much owing to the realist formula, is capable of explaining meta-ethical issues concerning moral disagreement, reason and motivation. Partly due to its explanatory capacity in relation to moral disagreement and motivation, it is possible for realism to counter significant arguments put forward in support of non-cognitivism and error-theory. In combination with the difficulties of non-cognitivism and error-theory mentioned in the introductory chapter, this indicates that realism is to be preferred to these views. Moreover, since realism, again much thanks to the realist formula, offers better explanations of the mentioned meta-ethical issues than reductionism, there is reason to believe that realism is to be preferred also to this view.

As is evident from this short summary of the reasoning thus far, the notion that objects have moral properties in virtue of having non-moral properties—particularly as it is expressed in the realist formula—is important for the arguments put forward in support of realism in the previous chapters. For this reason, I will in the present chapter examine how this dependence relation between moral and non-moral properties should be understood on realism. In particular, I will discuss some significant objections that can be directed against realism with regard to this relation.

In the next section, I call attention to the importance of realism and other meta-ethical views being capable of explaining the dependence of

moral properties on non-moral properties. In section 3, I argue that a traditional attempt to account for this relation is insufficient for realist purposes. In section 4, I return to the realist formula and argue that it does not have these difficulties. I also consider the objection that the realist formula is incompatible with particularism. In section 5, I introduce the well-known distinction between strong and weak supervenience and find that the realist formula implies both. I then argue in section 6 that in order to capture the mentioned dependence relation, strong supervenience is required. In sections 7 and 8, I discuss two arguments, proposed by Simon Blackburn and Jaegwon Kim, respectively, according to which strong and weak supervenience poses dilemmas for realism. I argue that there are reasons to believe that realism is able to avoid these dilemmas and abide by strong supervenience. However, in relation to Kim's argument, it is found that realism faces problems that I cannot deal with satisfactorily in the present thesis. In section 9, I consider J. L. Mackie's well-known arguments from queerness and argue that there are reasons to believe that realism is able to avoid them. However, it is found that Mackie's arguments raise issues that require further examination.

## *2. The Importance of Moral Dependence*

Previously I pointed out that one of our fundamental convictions as regards morality is that objects have moral properties in virtue of having non-moral properties, or to put the same point in the way I will make use of here, that moral properties *depend* on non-moral properties. (Above I reserved the term 'make' to signify the irreducible dependence relation between moral and non-moral properties that holds according to the realist formula.) The fundamental character of this notion is indicated by the fact that it seems to be a condition for being linguistically competent in respect of moral terms to know that these terms are to be applied in such a way that this dependence relation is recognised. Suppose, for example, that someone makes statements which indicate that she does not believe that a person's goodness depends on certain of the person's non-moral properties. She then seems committed to the view that it would be correct to claim that a person is good even if the person has no non-moral properties at all or that two persons can differ as regards goodness in spite of the fact that they do not differ in any of their non-moral properties. We would presumably regard

her statements as a sign that she is not linguistically competent in respect of 'good'. As I suggested earlier, since knowing that this necessary dependence relation should be recognised in one's application of moral terms arguably is a condition for being competent in the indicated respect, there is reason to believe that it is analytically necessary that moral properties depend on non-moral properties.

It is sometimes assumed that it is part of the mentioned competence to know that moral terms are to be applied in such a way that it is recognised that moral properties depend on *natural* properties, not non-moral properties.<sup>1</sup> According to this view, it might consequently be claimed that it is analytically necessary that moral properties depend on natural ones. However, I think this view is misguided. There are various conceptions of natural properties, and this view might perhaps be more plausible on some of these conceptions than on others. But consider a person who believes that a certain moral property depends on properties that are incontrovertibly non-natural. For example, think of a person who holds the view that an action's rightness depends on it having the property of being commanded by God.<sup>2</sup> Although most of us would deny this view, I do not think we would claim that the person lacks in linguistic competence with respect to 'right'. Hence, it does not seem analytically necessary that moral properties depend on natural properties.

However, it is reasonable to argue that the linguistic competence at issue involves other aspects concerning the relation between moral and non-moral properties than the one just mentioned. I have two such aspects in mind, but there might be others as well.

First, it can be argued that it is part of this competence to know that statements to the effect that an object has a certain moral property are to be justified by citing some of the object's non-moral properties. For instance, someone who is asked about the reason why she claims that a certain person is good, and who does not recognise that she is expected to justify her claim by citing some of the person's non-moral properties, may be suspected not to be fully linguistically competent in respect of 'good'.

Second, it can be argued that it is part of this competence to know

---

<sup>1</sup> See e.g. Blackburn (1993 (1985)), pp. 145–146; Dreier (1992), p. 15, and Smith (1994), p. 40.

<sup>2</sup> For a similar argument directed against Blackburn, see Klagge (1984), pp. 374–375.

that moral explanations are to be justified in terms of non-moral properties.<sup>3</sup> For instance, suppose someone claims that a person's goodness explains that she performed certain actions. The person who proposed the explanation may then be urged to justify this explanation. If she does not recognise that she is expected to provide the justification in terms of some of the assumed good person's non-moral properties, we may take this as an indication of her lacking linguistic competence with respect to 'good'. (I will return to a related observation in the next chapter.)

Thus, there are indications that the linguistic competence with regard to moral terms that concerns the relation between moral properties and non-moral properties involves more aspects than the notion that moral properties depend on non-moral properties. However, these additional aspects seem explainable in terms of the competence with regard to the dependence relation. As far as the epistemic aspect is concerned, when a person justifies a moral claim in terms of non-moral properties, this could be understood in terms of her trying to point at the non-moral properties on which the moral property in question depends. Similarly, as far as the explanatory aspect is concerned, when a person justifies her claim that an object's moral property explains that something has happened in terms of the object's non-moral properties, this could also be understood as an attempt to point to the non-moral properties on which the moral property depends. The fact that linguistic competence with regard to the dependence relation between moral and non-moral properties can be employed to explain other aspects of this competence accentuates the fundamental nature of this relation.

Another way to see the importance of the notion that moral properties depend on non-moral properties is to compare what holds for moral properties with what holds for certain other kinds of properties. In similarity with moral properties, it is claimed that these properties stand in a relation of dependence to some underlying properties. And in similarity with moral properties, it is sometimes claimed that these properties are irreducible to the underlying properties. However, in contrast to moral properties, this dependence relation is not connected in the same way to linguistic competence with respect to the discourse in question. One

---

<sup>3</sup> Cf. Audi (1993), pp. 101–103.

example of the latter phenomenon is provided by mental properties.<sup>4</sup> Suppose we understand Cartesian dualism as the view that there are two substances, the mental and the physical, which have separate existences. This view may be taken to imply that mental properties do not depend on physical ones. Although this position is implausible, it is conceivable. It seems at least possible to imagine that the fact that a person entertains a certain belief or harbours certain feelings is not dependent on her being in a certain physical state. This suggests that acknowledging the dependence of mental properties on physical properties is not part of the linguistic competence with respect to mental terms. This view finds further support in the observation that although we do not believe in the existence of disembodied beings, as for example angels or ghosts, we do not charge people who do so with not being linguistically competent with respect to mental terms. Colours provide another relevant example. It is claimed about colours that they depend on physical properties but are not reducible to such properties. And it does not seem to be part of being linguistically competent with respect to colour terms to recognise that colours depend on physical properties.<sup>5</sup>

The difference between mental properties and moral properties implies that a moral analogue to Cartesian dualism, moral dualism, according to which moral properties do not depend on any non-moral properties, is implausible. As far as I know, nobody has ever embraced this view, although meta-ethicists who believe that moral properties are irreducible have been accused of doing so.<sup>6</sup>

These remarks on the significance of the dependence of moral properties on non-moral properties suggest that any reasonable meta-ethical theory should be able to account for it. However, how this notion should be understood is open to dispute. This is especially so since the dependence relation might be understood more or less literally.

Meta-ethical views which claim that certain moral sentences involve moral terms in virtue of which moral properties are ascribed to objects and that moral properties exist so that some of these sentences are true,

---

<sup>4</sup> Cf. Blackburn (1993 (1985)), pp. 139–141; Jackson (1998), p. 119, and Klagge (1988), p. 466.

<sup>5</sup> Cf. Blackburn (1993 (1985)), pp. 144–145.

<sup>6</sup> See Warnock (1967), p. 14. For defences of Moore, see Baldwin (1985), pp. 24–30, and Cox (1970), 265–269.



characterise the dependence relation in terms of an ontological relation holding between moral and non-moral properties. This is presumably the most literal manner to understand the notion that moral properties depend on non-moral ones. Realists claim that moral properties ontologically depend on non-moral properties in a way that is incompatible with identity and hence with reduction. Reductionists, by contrast, claim that moral properties ontologically depend on non-moral properties in a way that entails identity. It might be tempting for realists to argue that we talk about the relation between moral and non-moral properties in a manner which indicates that the relation is asymmetric and hence not one of identity. Thus, we say that moral properties *depend* on non-moral properties, that non-moral properties *make* objects have moral properties, etc. However, we also seem to use this vocabulary when we are certain that identity is the case. I take it that we may say, for example, that the fact that a person is an unmarried man makes him a bachelor or that the fact that something is water depends on the fact that it is H<sub>2</sub>O. The fact that we use this asymmetric vocabulary in talking about the relation between moral and non-moral properties thus does not seem to be relevant to the dispute between realists and reductionists.

Meta-ethical views that reject the existence of moral properties cannot give the required explanation in terms of a straightforward ontological dependence relation. Error-theorists agree with realism and reductionism that certain moral sentences involve moral terms in virtue of which moral properties are attributed to objects, but deny that there are any moral properties and conclude therefore that all such sentences are false. Consequently, on this view there are no ontological dependence relations between moral and non-moral properties. (Indeed, as we will see below, one argument put forward in support of error-theory is that the existence of this dependence relation should be rejected because it would be metaphysically queer.) Non-cognitivists deny that moral sentences involve moral terms in virtue of which moral properties are ascribed to objects and hence that such sentences have truth-value. On this view, moral sentences do not ascribe any moral properties that depend on non-moral ones. Instead non-cognitivists have to account for the dependence relation in a non-ontological, and presumably less literal, way. Taking their point of departure in their view that moral sentences express non-cognitive states, advocates of

this view may argue that such states are to be had towards objects which have certain non-moral properties in a consistent way mirroring ontological dependence.<sup>7</sup>

In what follows, I will consider how the dependence of moral properties on non-moral properties should be accounted for according to realism.

## 2. *General Dependence*

A natural starting-point for a discussion of how realists should characterise the dependence of moral properties on non-moral properties is G. E. Moore's classical formulation of the relation he thinks holds between intrinsic value and intrinsic properties. This is especially so since one of Moore's purposes seems to be spell out his view that even though intrinsic value depends on intrinsic properties, it is not reducible to such properties.<sup>8</sup> Moore writes:

[I]f a given thing possesses any kind of intrinsic value in a certain degree, then not only must that same thing possess it, under all circumstances, in the same degree, but also anything *exactly like it* [as regards intrinsic properties], must, under all circumstances, possess it in exactly the same degree. Or to put it in the corresponding negative form: It is *impossible* that of two exactly similar things one should possess it and the other not, or that one should possess it in one degree, and the other in a different one.

I think this [. . .] proposition also is naturally conveyed by saying that the kind of value in question depends solely on the intrinsic nature of what possesses it.<sup>9</sup>

Since Moore is interested in intrinsic value, he assumes that this property depends on intrinsic properties, i.e., roughly, non-relational properties.<sup>10</sup> Realists, on the other hand, should not restrict the type of non-moral properties that moral properties can depend on in this way, but should allow for the possibility that they depend on 'extrinsic', i.e., roughly, relational

---

<sup>7</sup> See e.g. Blackburn (1993 (1971)), pp. 122, 125–126, and Hare (1952), pp. 133–136, 145–147, 159.

<sup>8</sup> See e.g. Moore (1993 (1922)), pp. 284, 295–297.

<sup>9</sup> Moore (1993 (1922)), p. 287. Other philosophers who also believe that moral properties or evaluative properties are irreducible have made similar claims. For two early examples, see Ross (1988 (1930)), pp. 116–123, and Sidgwick (1981 (1907)), pp. 208–209. Hare seems to be the one who made the dependence relation known as 'supervenience'; see e.g. Hare (1952), pp. 130–131, 153–155, and Hare (1989 (1984)), pp. 66–81. Jaegwon Kim has developed various concepts of supervenience, see e.g. Kim (1993 (1984)), pp. 53–78, and Kim (1993 (1990)), pp. 131–160. (I will return to Kim's views on supervenience later on in this chapter.)

<sup>10</sup> It might be argued that this does not even hold for intrinsic value; see e.g. Rabinowicz and Rønnow-Rasmussen (2000), pp. 33–49.

non-moral properties.

If we map over what Moore says about the relation between moral properties and non-moral properties and take into consideration what was said in the last section about it being analytically necessary that moral properties depend on non-moral properties, we obtain the following claim:

*The general formula:* It is analytically necessary that, for any objects  $x$  and  $y$ , if  $x$  and  $y$  are exactly similar as regards their non-moral properties, then  $x$  and  $y$  are exactly similar as regards a moral property  $M$ , i.e. both  $x$  and  $y$  have  $M$  or neither has.

This formula can be said to capture a general dependence of moral properties on non-moral properties, since it expresses the idea that moral properties depend on the category or family of non-moral properties as a whole, rather than on particular subsets of such properties.

It is worth noticing that the general formula is a non-symmetric claim in a certain sense: that two objects are exactly similar as regards their non-moral properties entails that they are exactly similar as regards a moral property, but the converse does not follow. But while it is non-symmetric, it is not asymmetric either; it is neither nor. Although it does not entail that sameness as regards a moral property is followed by sameness as regards non-moral properties, a claim which would be quite implausible, it is not incompatible with this being the case. However, it would be easy to rewrite the formula so as to exclude this possibility and hence make into a kind of asymmetric claim.

At least three considerations indicate that realists should not stay satisfied with the general formula.

Firstly, it seems fairly vacuous. For example, it might be argued that two objects cannot be exactly similar as regards all their non-moral properties because they have to differ as regards at least some relational non-moral properties. But in that case it seems of little significance that the fact that two objects are exactly similar as regards all their non-moral properties entails that they are similar as regards a moral property. In particular, it is difficult to see that this would say anything important about the dependence of moral properties on non-moral properties.<sup>11</sup>

---

<sup>11</sup> There are related objections that can be directed against the general formula. For example, that two objects are exactly similar as regards their non-moral properties is compatible with the fact that they do not have any non-moral properties at all, in which case it seems mistaken to take the general formula to say that moral properties depend on

Secondly, it is compatible with moral properties being identical with non-moral properties. The general formula says that sameness as regards non-moral properties entails sameness as regards a moral property. This can be the case even if a moral property is identical with a non-moral property. Because realists deny this view, they should not stay satisfied with the general formula.

There is a third reason why realists should not stay satisfied with the general formula. To see this, consider that there are non-moral properties that are irrelevant to whether objects have a certain moral property.<sup>12</sup> For instance, a person's non-moral property of having a certain weight does not seem relevant to her being morally good. Still, such properties are included among the non-moral properties on which moral properties, according to the general formula, depend. This may perhaps not seem to be a difficulty for this claim, since it should be understood as an attempt to capture the idea that moral properties depend on the category of non-moral properties as such, the whole family of non-moral properties, not on any particular non-moral properties. However, we also take moral properties to depend on particular subsets of non-moral properties, where these subsets only contain non-moral properties that are relevant to objects having the moral properties. The general formula has thus to be supplemented with a dependence claim that captures this idea.

The last difficulty could perhaps be explained by appeal to linguistic competence concerning moral terms. It seems to be a part of that competence not only to know that moral terms are to be applied in such a way that it is recognised that moral properties depend on non-moral ones, but also in such a way that it is recognised that they depend on particular subsets of non-moral properties. Consider someone who makes statements which indicate that she believes that a person's goodness depends on all of the person's non-moral properties. It seems plausible to say that we would take this as an indication of her not being linguistically competent with

---

non-moral properties. (Cf. Bailey (1998/99), p. 55, and Grimes (1991), pp. 82, 85.) To this objection it might perhaps be responded that it is impossible for an object not to have any non-moral properties. For example, it might be argued that all objects have various relational non-moral properties. The objection would perhaps be more troublesome on the conception that moral properties necessarily depend on natural properties, since it might not be plausible to rule out the possibility that objects entirely lack natural properties.

<sup>12</sup> Cf. Dancy (1981), p. 374; Griffin (1992), pp. 314–316; Grimes (1991), pp. 88–89, and Kim (1993 (1984)), p. 66.

respect to ‘good’, since she does not seem to understand which kind of non-moral properties that is relevant and thus could make a difference for whether the term applies. If linguistic competence and the relevance of non-moral properties are related in this way, it could explain why we find the idea that all non-moral properties are relevant so peculiar.

#### 4. *Realist Dependence*

In the last section, I argued that for at least three reasons, the general formula is unsatisfactory as a formulation of a realist account of the dependence of moral properties on non-moral properties. This gives us reason to return to the realist formula and inquire whether it avoids these difficulties. Recall the realist formula as it was stated in chapter 4:

*The realist formula:* (i) It is analytically necessary that, for any object  $x$ , and for any moral property  $M$ , if  $x$  has  $M$ , then there is a set of non-moral properties  $G$  such that (A)  $x$  has  $G$ , and (B) it is synthetically necessary that, for any object  $y$ , if  $y$  has  $G$ , then  $y$  has  $M$ . (ii)  $M$  is not identical with any non-moral property.<sup>13</sup>

It seems evident that the realist formula avoids the two first difficulties of the general formula. First, it cannot be accused of being vacuous, among other things because it, unlike the general formula, does not refer to all non-moral properties of the objects in question. Second, in virtue of including (ii), the realist formula is, unlike the general formula, incompatible with a moral property being identical with a non-moral property. The realist formula also seems to avoid the third problem with the general formula. The kind of set of non-moral properties it refers to does not need to contain all non-moral properties of an object. On the contrary, it is plausible to assume that such a set only consists of a subset of an object’s non-moral properties. It is therefore plausible to claim that the realist formula is able to give expression to the notion that an object’s moral property depends only on a set of non-moral properties that is relevant to the object having the moral property.

---

<sup>13</sup> In the last section, I argued that the kind of dependence Moore takes to hold between intrinsic value or goodness and intrinsic properties in ‘The Conception of Intrinsic Value’ is insufficient for purposes of realism. It is therefore interesting to note that when Moore reconsiders the matter in ‘Reply to my Critics’, he provides an account that comes close to the realist formula: ‘I have always supposed that it [i.e. goodness] did so “depend,” in the sense that, if a thing is good (in my sense), then that it is so *follows* from the fact that it possesses certain natural intrinsic properties, which are such that from the fact that it is good it does *not* follow conversely that it has those properties’ (Moore (1942), p. 588).

However, it might be objected that even if the kind of set of non-moral properties referred to in the realist formula may contain only what is relevant to the object having the moral property, the realist formula does not *guarantee* that this is the case. There is nothing in the realist formula that guarantees that the kind of set of non-moral property it refers to only contains what is relevant to an object having the moral property and hence is what its moral property depends on.

I would like to reply to this objection with two comments. We should first observe that to answer it in a way that the objector would find completely satisfactory, it would probably be necessary to construe a kind of set of non-moral properties that only contains relevant non-moral properties and that could be inserted in the realist formula.<sup>14</sup> However, I think it is quite difficult to do so in a manner that is neutral in the desirable way. In order not to beg any significant questions, the realist formula should be compatible with various views about what makes objects have a moral property. It should then leave open the possibility that such a set of non-moral properties contains merely one non-moral property, but also that it contains a conjunction of a number of such properties, properties which might be heterogeneous. Moreover, it should leave open the possibility that these properties can be related to each other in different ways. Granted the various ways such a set of non-moral properties could look like, it seems difficult to construe one abstract kind of set of non-moral properties that is neutral in the desirable way. Second, it might be replied that it is not a task of meta-ethics to construe such a set. As we have seen, the realist formula is able to give expression to the view that an object's moral property depends only on a set of non-moral properties that is relevant to the object having the moral property. There, it might be claimed, ends the task of meta-ethics. It is then up to a normative conception to specify what such a set of non-moral properties looks like.

In this connection, it is suitable to take the opportunity of considering a certain objection against the realist formula. It might be suspected that it is incompatible with a much debated view in contemporary ethics known as 'particularism'. This contention raises a number of issues that are impossible to do justice to here. However, I think it is reasonable to argue that the

---

<sup>14</sup> Jaegwon Kim proposes one quite general way to construe sets of relevant properties that can constitute the supervenience base of a supervenient property; see e.g. Kim (1993 (1984)), pp. 58–59.

realist formula is compatible with the main tenet of particularism. It is presumably irreconcilable with a view commonly associated with particularism, but it might be argued that this problem is not especially serious.

When Jonathan Dancy, the most prominent advocate of particularism, sets out to define this view, he writes:

Particularism, I want to say, is an expression of a general holism in the theory of reasons; it is the application of holism to the moral case. Holism in the theory of reasons holds that a feature that is a reason in favour in one case may be no reason at all in another, and in a third may even be a reason against.<sup>15</sup>

Thus understood, particularism claims that the relevance of non-moral properties is *context dependent*. As far as I understand, this view is generally considered to be the main tenet of particularism.<sup>16</sup> It can perhaps be formulated in the following way: a non-moral property which, when instantiated in one object, contributes to the object having a certain moral property, might, when instantiated in another object, contribute to that object *not* having the moral property, and might, when instantiated in yet another object, contribute in neither of these ways. Suppose A is such a property. The reason why A's relevance varies in this way is that some of the object's other non-moral properties determine whether A is relevant and, if it is, which relevance it has. The relevance of A is thus context dependent, where the context is made up by other non-moral properties of the object. Particularists find support for this view in various thought experiments of which the following might serve as an example.<sup>17</sup> Suppose an action causes pleasure and that we think that it is right because it has that non-moral property. This fact may have us believe that causing pleasure is a non-moral property which always contributes to actions being right. To see that this is not at all evident, imagine that the action of executing a person in public causes pleasure in the audience.<sup>18</sup> In that case, we might be inclined to say, the property of causing pleasure does not contribute to the action being right. Perhaps we might even be inclined to say that it

---

<sup>15</sup> Dancy (1999a), p. 144.

<sup>16</sup> Cf. Crisp (2000), p. 34; Dancy (1993), p. 60; Hooker (2000), p. 6; Kihlborn (2002), pp. 23–28; McNaughton (1988), pp. 193–194, and Raz (2000), p. 59. For a defence of the opposite of holism, atomism, see Alm (2004), pp. 312–331.

<sup>17</sup> For various examples, see e.g. Crisp (2000), pp. 36–37; Dancy (1993), pp. 60–62, and Sinnott-Armstrong (1999), pp. 3–4.

<sup>18</sup> For this example, see Dancy (1993), p. 61, and McNaughton (1988), p. 193.

contributes to the action *not* being right. If this is correct, the relevance of causing pleasure varies depending on the context made up by other non-moral properties, e.g. the property of being a public execution.

I think it can be suggested that the realist formula is compatible with the relevance of non-moral properties being context dependent. A set of non-moral properties that according to the realist formula makes objects have a moral property might consist of a number of non-moral properties. Within such a set, the relevance of a moral property might be context dependent, where the context is made up by the other non-moral properties in the set.

Consider the following simple example of how this might be possible. (There are presumably more sophisticated accounts, but here I am merely interested in the principal point that the realist formula is compatible with the relevance of non-moral properties being context dependent.<sup>19</sup>) Suppose that an object has the following set of non-moral properties: A & -B & C. Suppose further that this set of non-moral properties makes the object have M; that is, this set of non-moral properties is of the kind referred to in the realist formula. In this set, A might contribute to the object having M. This can perhaps be understood in the following way: in this set, A is such that if the object had not had A, it would not have had M. However, suppose another object has the following set of non-moral properties: A & B & C. Suppose further that this set does not make the object have M. In this set, A might contribute to the object not having M. More exactly, in this set, A is such that if the object had not had A, it would have had M. Or, as regards this set, A does not contribute in either way; that is, neither of the two counterfactuals holds. Thus, whether A is relevant, and, if it is, which relevance it has, is determined by the other non-moral properties in the respective set of non-moral properties, i.e. A's relevant context. In the sets at issue, the active part of the context is -B and B, respectively.

We may apply this picture of the context relevance of non-moral properties to the thought experiment mentioned above. Let the moral property, M, be 'rightness', A be 'causing pleasure' and B 'being a public execution'. In the first set of non-moral properties mentioned above (A & -B & C), A contributes to the object having M because it figures in a context

---

<sup>19</sup> One potential difficulty with the approach suggested here is that it makes use of complements of properties, 'negative properties'.



partly made up by -B. In the second set of non-moral properties mentioned above (A & B & C), A contributes to the object not having M, or does not contribute in either way, because it figures in a context made up partly by B.

The following objection might be directed against this suggestion. The kind of set of non-moral properties that according to the example above makes objects have a moral property consists of a conjunction of non-moral properties. Assume that conjunction is a proper way of forming properties so that a conjunction of properties itself is a property. In that case, such a set of non-moral properties is a non-moral property. Now, such a non-moral property appears to contain all non-moral properties that are relevant to a certain object having a moral property.<sup>20</sup> But this means that it does not have any context that determines its relevance. The relevance of such a non-moral property is in other words not context dependent. Hence, the proposal above is incompatible with the view that the relevance of non-moral properties is context dependent.

However, I do not think the suggestion above is incompatible with particularism; at least, I do not think it runs counter to the spirit of this view. The sort of non-moral properties particularists appeal to in their examples, and whose relevance is claimed to be context dependent, are simple properties, e.g. 'causing pleasure', 'breaking a promise', 'lying', 'being unkind' and 'being cruel'. However, the kind of non-moral properties that would make objects have a moral property if simple non-moral properties as these are context dependent would be complex, perhaps *very* complex, non-moral properties. As far as I see, it is not incompatible with particularism—at least not with the spirit of this view—to claim that such complex properties are not context dependent.

Thus far, I have argued that the realist formula is compatible with the main tenet of particularism, i.e. the view that the relevance of non-moral properties is context dependent. However, particularism is often associated with another view, namely with the claim that there are no true moral principles. This view does indeed seem incompatible with the realist formula. The reason is that the realist formula contains (**B**): a synthetically necessary implication from a set of non-moral properties to a moral property. Such an implication constitutes a kind of moral principle. Indeed,

---

<sup>20</sup> This will become clearer when we consider so-called enabling conditions below.

since there might be a number of sets of non-moral properties that make objects have a certain moral property, a number of moral principles of this kind might hold.

It is sometimes assumed that the context dependence of non-moral properties means that there are no true moral principles.<sup>21</sup> This assumption seems unfounded. As we have just seen, the context dependence of the relevance of (simple) non-moral properties is compatible with the realist formula. The realist formula involves (**B**), which, as just mentioned, is a kind of moral principle. Hence, the relevance of non-moral properties being context dependent seems compatible with the truth of at least one kind of moral principles.

The claim that there are no true moral principles finds instead support in another view advocated by Dancy: that there is a distinction between, on the one hand, the non-moral properties that *make* objects have a certain moral property and, on the other hand, the non-moral properties that merely *enable* other non-moral properties to make objects have the moral property, but which do not themselves have this function, what are called ‘enabling conditions’.<sup>22</sup> Dancy admits that he does not know how to draw this distinction in any clear way.<sup>23</sup> Consequently, it is unclear whether a certain non-moral property on this view should be classified as belonging to the first or second category. However, sets of non-moral properties of the kind described above, and which are hypothesised to make objects have a moral property if the relevance of non-moral properties is context dependent, may contain non-moral properties of a type that Dancy explicitly argues are enabling conditions. To see this, recall the set of non-moral properties mentioned above, A & -B & C, which might be assumed to make objects have a moral property M. In the thought experiment mentioned above, -B represents ‘not being a public execution’. As Dancy

---

<sup>21</sup> See e.g. Dancy (1993), p. 66.

<sup>22</sup> See e.g. Dancy (1993), pp. 22–26, 55–58; Dancy (1999a), pp. 148–150, and Dancy (1999b), pp. 26–29. When I discuss the objection issuing from enabling conditions, I follow Dancy in using ‘make’ in a wider sense, not connected exclusively to the realist formula.

<sup>23</sup> Dancy (1999a), p. 148. Dancy admits this in a response to an objection raised by Kasper Lippert-Rasmussen: Lippert-Rasmussen (1999), pp. 101–104. Dancy’s admission concerns his wider tripartite distinction between ‘foreground properties’, ‘active background properties’ and ‘inert background’ properties, but as far as I understand, it applies also to the simpler distinction I focus on here. Dancy’s view on enabling conditions is criticised by e.g. Kihlborn (2002), pp. 62–65; Lippert-Rasmussen (1999), pp. 99–104; Raz (2000), pp. 68–69, and Sinnott-Armstrong (1999), pp. 2–8.

sees matters, this non-moral property enables A, i.e. ‘causing pleasure’, to make the action in question have M, i.e. right, but it is not itself a right-making property.<sup>24</sup> Dancy writes:

The action’s not having a property strong enough to make it right is not a property in virtue of which it is wrong, though it is something required of it if it is to be wrong. Not having a countervailing property is something without which it would be right and with which it is bound to be wrong; but it is not what makes it wrong—that is done by more ordinary things like being unkind or cruel.<sup>25</sup>

In Dancy’s view, non-moral properties of the mentioned type are thus enabling conditions and not among the non-moral properties which make objects have a certain moral property. Dancy’s argument for this view, when applied to rightness, is that they are not ‘something for which we judge the action to be right’.<sup>26</sup> Generally put, the argument is that such non-moral properties are not among our reasons for judging that objects have a moral property.

We can now see that Dancy’s distinction between non-moral properties that make objects have a certain moral property and enabling conditions may be thought to have at least two consequences for the realist formula. First, it can be taken to imply that the realist formula may not state a set of non-moral properties that makes objects have a moral property. According to the mentioned distinction, a set of non-moral properties of the kind referred to in the realist formula may contain a non-moral property that is not among the properties that make objects have a moral property, but only enables other properties to do so. (For example, it might involve a property as -B.) As a result, it may seem mistaken to claim that such a set of non-moral properties makes objects have a moral property. Second, the distinction can be taken to imply that moral principles of the kind stated in **(B)** may turn out to be false. Assume that true moral principles, if there are any, state non-moral properties that make objects have a moral property.<sup>27</sup> As we have seen, the mentioned distinction suggests that in case a set of

---

<sup>24</sup> It may be assumed that a property of the kind ‘being a public execution’ is such that its absence is necessary for an object to have a certain moral property; such properties might be called ‘defeaters’. For a classification for various types of defeaters, see Sinnott-Armstrong (1999), p. 5.

<sup>25</sup> Dancy (1993), p. 77. See also e.g. Dancy (1993), pp. 80–81, and Dancy (1999b), p. 26.

<sup>26</sup> Dancy (1993), p. 81.

<sup>27</sup> Cf. Dancy (1993), pp. 76–77, and Dancy (1999b), p. 26.

non-moral properties contains an enabling condition, it fails to make objects have a moral property. It follows that a version of **(B)** which refers to such a set of non-moral properties does not constitute a true moral principle. Since the second consequence can be seen as a special case of the first consequence, I will focus on the first one.

I am not quite sure that this line of reasoning can be met in a completely satisfactory way. However, I would like to respond with the following observations.

First, it should be kept in mind that it is on the assumption that non-moral properties are context dependent in the way particularists claim that these two consequences arise. According to this view, the contribution a non-moral property makes to an object's moral property depends on its context. This implies that it is not the case that, for any object, if it has such a non-moral property, it has a certain moral property. For example, it is not the case that, for any object, if it has A, it has M. This means that such a property alone cannot constitute a set of non-moral properties of the kind referred to in the realist formula. On the contrary, for a set of non-moral properties to be such that, for any object, if it has that set of non-moral properties, it has a moral property, we have to add what Dancy considers to be enabling conditions to the set. For example, we have to add properties such as  $\neg B$  to a set of non-moral properties involving A for such an implication to hold. It is only if a set of non-moral properties is supplemented in this way that it can figure in the realist formula. As we have seen, according to Dancy's distinction, it might be argued that such a set of non-moral properties does not make objects have a moral property. However, assume that the relevance of non-moral properties is not context dependent. In that case, there might be a non-moral property, or a combination of non-moral properties, such that, for any object, if it has that non-moral property or combination of non-moral properties, it has a moral property. Such a non-moral property, or combination of non-moral properties, could then alone constitute a set of non-moral properties of the kind referred to in the realist formula.<sup>28</sup> On this alternative, there is no need to add any enabling conditions. Such a set of non-moral properties could then be what makes objects have a moral property.

---

<sup>28</sup> Granted, of course, that the implication is synthetically necessary.

Second, it seems reasonable to suggest the adoption of a more technical sense of ‘make’, according to which also non-moral properties that Dancy’s classifies as enabling conditions are among the properties that make objects have a certain moral property. This means that even if a set of non-moral property of the kind referred to in the realist formula involves a property that Dancy classifies as an enabling condition, it can be such that it makes objects have a moral property. However, it would do so in a sense that presumably diverges somewhat from the ordinary sense of ‘make’. As far as I understand, this adjustment in our usage of the term would be rather innocent, since little of real philosophical importance seems to follow from it. This contention is reinforced by the fact, admitted by Dancy, that there does not seem to be any clear distinction between non-moral properties that make objects have a moral property and enabling conditions. It should also be pointed out that Dancy does not deny that there are true implications that take us from a set of non-moral properties to a moral property.<sup>29</sup> What he stresses is that such an implication does not state non-moral properties that make objects have a moral property and that it therefore does not constitute a true moral principle. This suggests that this aspect of the debate about particularism—unlike the aspect that concerns context dependence—can be understood to concern how the term ‘make’ should be used, rather than some more fundamental philosophical issue.

Third, it should be observed that the realist formula indirectly involves a set of non-moral properties that does not contain any non-moral properties that Dancy classifies as enabling conditions. Such a set of non-moral properties is a subclass of a set of non-moral properties of the kind referred to in the realist formula. A set of that kind contains only properties that make objects have a certain moral property, in the stricter sense of ‘make’ Dancy appeals to. This being the case, it might be maintained that the realist formula, after all, succeeds in stating a set of non-moral properties that—in the stricter sense of ‘make’—makes objects have a moral property.

Lastly, we may question Dancy’s view that the non-moral properties that he classifies as enabling conditions are not among our reasons for judging that objects have a certain moral property. Joseph Raz argues, in a critical discussion of particularism, that we should distinguish between the reasons people cite in support for their moral judgements and the reasons

---

<sup>29</sup> See e.g. Dancy (1999b), p. 26.

there actually are for these judgements.<sup>30</sup> These may of course come apart in different ways and for different reasons. However, what is important in the present context is that there might be features that people do not cite as reasons for their moral judgements even though they actually constitute such reasons. For example, suppose we ask someone about her reasons for asserting that a certain person is good. She is likely to mention features of the person that she takes to be in some way characteristic of the person being good, features she experiences as somehow ‘standing out’ and being especially important in consideration of the person’s goodness. However, there may be other features that are less conspicuous and lie more in the ‘background’ which also constitute reasons to judge that the person is good, but which she does not mention. It seems reasonable to suppose that what Dancy classifies as enabling conditions are among these features. For example, although a person would not mention ‘not being cruel’ among the non-moral properties that make a person good, this does not mean that this non-moral property does not comprise a reason to judge that the person is good. Moreover, that the non-moral properties Dancy classifies as enabling conditions may be among the reasons to judge that something has a moral property might be something people would agree to, if we were to call their attention to these properties.<sup>31</sup> Suppose, for example, that the person mentioned above claims that someone is good because she is friendly and benevolent. We may point out to her that if the person had been cruel, she would not have been good. Suppose we then ask if this does not mean that ‘not being cruel’ is among the features that constitute reasons to judge that the person is good. It seems reasonable to assume that she would agree.

If the last comment is plausible, we are in the position to question Dancy’s view that the non-moral properties that he classifies as enabling conditions are not among our reasons for judging that objects have a certain moral property. This suggests that his argument for claiming that these non-moral properties are enabling conditions and not among the non-moral properties that make objects have a moral property is mistaken. In other words, we may question Dancy’s distinction. This, in turn, would mean that the two consequences described above do not follow. However, it should be admitted that I have not demonstrated that these considerations apply to

---

<sup>30</sup> Raz (2000), pp. 61–70. Cf. Lippert-Rasmussen (1999), p. 102.

<sup>31</sup> Cf. Raz (2000), p. 67.

all non-moral properties Dancy classifies as enabling conditions. There might be non-moral properties that are so outlandish that they cannot be assumed to belong to the non-moral properties that make objects have a moral property, but have to be classified as enabling conditions.<sup>32</sup>

### 5. *The Realist Formula and Supervenience*

It is now time to call attention to something that has remained implicit so far: the close resemblance between the realist formula and supervenience claims. Like the realist formula, supervenience claims are generally intended to capture a dependence relation holding between properties, e.g. the dependence of moral properties on non-moral properties. We may distinguish between two supervenience claims that concern the relation between these two types of properties.<sup>33</sup> Consider first:

*Weak supervenience:* It is necessary that, for any object *x*, and for any moral property *M*, if *x* has *M*, then there is a set of non-moral properties *G* such that **(a)** *x* has *G*, and **(b)** for any object *y*, if *y* has *G*, then *y* has *M*.

The occurrence of ‘necessary’ that opens this claim binds the formula as a whole. However, the implication in **(b)** is not prefixed with ‘necessary’ and consequently does not extend to all possible worlds. In weak supervenience, **(b)** merely says that any object *within* a possible world which has a set of non-moral properties *G* has *M*. It does not say that any object in *all other* possible worlds which has *G* has *M*. The correlation between *G* and *M* does not, so to speak, spread to other possible worlds. Consider next:

*Strong supervenience:* It is necessary that, for any object *x*, and for any moral property *M*, if *x* has *M*, then there is a set of non-moral properties *G* such that **(a)** *x* has *G*, and **(b)** it is necessary that, for any object *y*, if *y* has *G*, then *y* has *M*.

As before, the first occurrence of ‘necessary’ binds the formula as a whole. However, in strong supervenience the second occurrence of ‘necessary’ binds the implication in **(b)**. Since the implication in **(b)** now is prefixed in this way, it says that it holds in all possible worlds that any object which has *G* has *M*. Thus, the correlation between *G* and *M* does spread to other

---

<sup>32</sup> The non-moral property -B above, ‘not being a public execution’, is perhaps a case in point.

<sup>33</sup> As far as I know, Jaegwon Kim was the first one who observed the difference between strong and weak supervenience; see e.g. Kim (1993 (1984)), pp. 57–67.

possible worlds.

The realist formula and the two supervenience claims differ in two important respects.

First, whereas the realist formula, in virtue of involving the condition (ii), rules out that a moral property is identical to a non-moral property, the two supervenience claims are neutral in this regard. They lack (ii) and are consequently compatible both with that identity holds and that it does not. The fact that they are neutral in this respect suggests that they might be adopted by realists and reductionists alike.

Second, whereas the realist formula specifies the kind of necessities at issue, the two supervenience claims do not. They are consequently neutral also in this respect. This means that various versions of weak and strong supervenience can be generated by the insertion of different kinds of necessities in each respective claim.

Moreover, it is important to notice that, on the appropriate reading of the two occurrences of ‘necessary’, the realist formula entails strong supervenience. Of course, it also entails weak supervenience. As we will see in sections 7 and 8, arguments have been proposed to the effect that realists should not adopt strong supervenience. If these arguments are correct, they would also show that realists should not adopt the realist formula. Since these arguments have been formulated in terms of strong supervenience, I will discuss them in terms of this claim rather than in terms of the realist formula.

## 6. *Strong or Weak Supervenience?*

It may now be asked whether realists should adopt strong supervenience or weak supervenience.<sup>34</sup> The answer seems to be that realists should opt for strong supervenience. The primary reason is that weak supervenience is too weak to capture the notion that an object’s moral property depends on a set of its non-moral properties whereas strong supervenience succeeds in doing so. This reason is quite general and applies to other properties as well. Among others Jaegwon Kim has argued that in order for a supervenience claim to state a dependence relation between properties, it cannot be understood as weak supervenience; it has to be understood as strong

---

<sup>34</sup> In chapter 4, I argued that realists should adopt the realist formula, which entails strong supervenience. Here I will for the moment ignore this result and consider whether there are any other reasons for realists to adopt strong supervenience.



supervenience.<sup>35</sup>

To see why this is the case, consider the following example. Suppose we want to claim that a person's goodness depends on her having a set of non-moral properties G. As we saw in the last section, weak supervenience requires merely that *within* a possible world any person who has G is good. This means that weak supervenience is compatible with that people in another world who have G are not good. This indicates that, on weak supervenience, it would be mistaken to claim that a person's goodness depends on her having G, since a person *could* have this set of non-moral properties and yet not be good. If a person could have G and yet not be good, we would not consider the co-instantiation of goodness and G as a matter of dependence, but rather as a matter of coincidence: that people who have G *happen* to be good. In Kim's words: 'Determination or dependence is naturally thought of as carrying a certain modal force: if being a good man is dependent on, or is determined by, certain traits of character, then having these traits must *insure* or *guarantee* being a good man'.<sup>36</sup> Strong supervenience provides the required supplement. Applied to the example above, it says that it holds in all possible worlds that any person who has G is good. This indicates that it contains the modal strength that is required for the relevant dependence relation to hold. The same reasoning applies to other moral properties. Hence, realists should adopt strong supervenience rather than weak.

Furthermore, it should be born in mind that much of our moral thinking is constituted by thought experiments. Generally put, in thought experiments we ask how things would be under certain specified conditions by imagining how things are in a certain possible world. For instance, being convinced that envy is relevant to whether a person is good, I might ask if I would be a better person were I not so envious by imagining how I am in a possible world in which I am less envious than I am in the actual world. We often trust the results of such thought experiments and draw certain conclusions from them; for one thing, we let our moral decisions be guided by them. For example, convinced by the result in my thought experiment—that I would be a better person were I less envious—I might

---

<sup>35</sup> See e.g. Kim (1993 (1984)), pp. 59–61, and Kim (1993 (1990)), pp. 143–144. In relation to morality, similar points have been made by e.g. Blackburn (1993 (1985)), p. 132, and Depaul (1987), pp. 431–432.

<sup>36</sup> Kim (1993 (1984)), p. 60.

decide to try to get less envious. However, if only weak supervenience were the case, thought experiments would not be of any help, since we then would not be justified to hold beliefs about one possible world that are based on what we believe about another possible world.<sup>37</sup> In particular, we would not be justified to believe that non-moral properties contribute, or fail to contribute, to objects having moral properties in the same way in our world as they do in the possible worlds employed in our thought experiments. However, since thought experiments often seem reliable, there is reason to believe that weak supervenience is insufficient. Strong supervenience allows us to trust them, and there is therefore reason for realists to adopt this claim.

### *7. Blackburn: Strong or Weak Supervenience—a Dilemma*

According to a much discussed argument put forward by Simon Blackburn, realism faces a dilemma with regard to strong and weak supervenience. Blackburn's argument displays, roughly put, the following structure: **(i)** Realists have reason to prefer strong supervenience to weak. **(ii)** However, strong supervenience results in reduction. **(iii)** Realists are then advised to adopt weak supervenience. But then they face what Blackburn calls 'the explanatory problem' which they fail to account for. Thus, realism faces a dilemma: strong supervenience leads to reduction whereas weak supervenience leads to the explanatory problem. **(iv)** A certain version of non-cognitivism can account for the explanatory problem and should thus be preferred.<sup>38</sup> (Here I will only be concerned with steps **(i)**–**(iii)**.)

I think it is reasonable to argue that realism can avoid the dilemma presented by Blackburn. However, a caveat is in place: I find Blackburn's argument difficult to understand, so what I argue against is to a great extent an interpretation of what he says.

Blackburn does not characterise supervenience in the terminology adopted above. However, I have chosen to reformulate Blackburn's

---

<sup>37</sup> Cf. Kim (1993 (1984)), pp. 60–61.

<sup>38</sup> The argument is presented in Blackburn (1993 (1971)), pp. 111–129; Blackburn (1984), pp. 182–187, and Blackburn (1993 (1985)), pp. 130–148. Here I will mainly follow Blackburn's latest formulation of the argument. In trying to understand it, I have been helped by the various critical comments it has given rise to: Bovens and Draï (1999), pp. 241–245; Brueckner (2002), pp. 67–70; Dreier (1992), pp. 13–38; Elliot (1987), pp. 133–137; Klagge (1984), pp. 370–380; Klagge (1987), pp. 312–315; McFetridge (1985), pp. 245–258; Noonan (1985), pp. 78–85; Shafer-Landau (2003), pp. 84–89; Yassenchuk (1995), pp. 84–89, and Zangwill (1995), pp. 240–262.

supervenience claims to this terminology so as to make it easier to determine the bearing of his arguments on realism.<sup>39</sup> Despite this, I will in this section make use of some of Blackburn's names of certain theses.

Let us now consider the relevant steps in Blackburn's argument.

(i) Blackburn concedes that strong supervenience appears to be what realists should opt for. The reason, which Blackburn apparently does not regard as conclusive, is similar to the one I mentioned above: on strong supervenience, but not on weak, is it reasonable to hold that an object's moral property depends on a set of its non-moral properties or, as Blackburn puts it, that a set of the object's non-moral properties 'underlies' its moral property.<sup>40</sup>

(ii) Blackburn suggests, however, that there is reason for realists not to adopt strong supervenience. In the terminology adopted here, he argues as follows. Assume that there is an object which has a moral property M. On strong supervenience, it then follows that there is a set of non-moral properties G such that the object has G and it holds necessarily that whatever object has G has M. That is, a necessary implication of the following kind holds:

---

<sup>39</sup> Blackburn explains weak supervenience as the thesis 'that as a matter of necessity, if something x is F, and G\* underlies this, then anything else in the physical or natural (or whatever) state G\* is F as well'. Formally Blackburn represents weak supervenience thus:

(S)  $N((\exists x)(Fx \& G^*x \& (G^*x \cup Fx)) \supset (y)(G^*y \supset Fy))$

'N' should be read 'necessarily'; 'Fx' is a particular moral truth or fact; 'G\*x' is a 'definite total set of G truths', a set of natural truths which do not include all the truths of this kind about the object. 'U' represents 'the relation [. . .] that holds when one "underlies" the other'. Blackburn represents strong supervenience thus:

(?)  $N((\exists x)(Fx \& G^*x \& (G^*x \cup Fx)) \supset N(y)(G^*y \supset Fy))$ .

(Blackburn (1985 (1993), pp. 131–133.)

As can be seen, Blackburn makes use of the term 'underlie' and its formal equivalent 'U' in his account of supervenience. One difficulty with this notion is that if G\* underlies F, it seems reasonable to assume that F depends on G\*, in which case the 'underlying relation' is nothing but the converse of the supervenience relation in so far as this is understood to express dependence. This means, however, that one cannot use 'underlie' or 'U' in an explanation of what supervenience amounts to, unless one is prepared to accept that the account is circular. On the other hand, if one understands 'underlie' in a way which does not mean that it is related to dependence in the indicated way, it is difficult to see which relevance it has for an explication of supervenience. It can further be argued that on Blackburn's characterisation of supervenience, weak supervenience (S) becomes a rather peculiar claim. If, as the first antecedent of (S) says, G\* underlies F it seems hard to deny that each object in any possible world which has G\* also has F, since a reasonable interpretation of 'underlie' means that F depends on G\*, and this relation presumably holds across possible worlds, as we saw in the last section.

<sup>40</sup> Blackburn (1993 (1985)), p. 132. Blackburn believes that supervenience claims are intended to state that moral properties depend on natural properties rather than on non-moral properties. However, this difference between our formulations of supervenience is not important in the present context.

(**N**) It is necessary that, for any object *x*, if *x* has *G*, then *x* has *M*.<sup>41</sup>

Now, according to Blackburn, the truth of an implication like (**N**) means that *M* has been reduced. To avoid this conclusion, Blackburn believes that realists should deny (**N**). To make this explicit, they should assert the following kind of claim:

(**P**) It is possible that there is an object *x* that has *G* but which does not have *M*.

(iii) In Blackburn's view, in order to account for the notion that an object's moral property depends on a set of its non-moral properties, realists should adopt weak supervenience. That is, they should maintain that it holds in any possible world that if there is an object which has a moral property *M*, it has a set of non-moral properties *G*, and then, within a possible world where an object has *G*, every object which has *G* has *M*. Moreover, in order to avoid reduction, realists should combine weak supervenience with (**P**). As a result, they maintain that there is a possible world in which objects that have *G* do not have *M*. On this view, the relation between *G* and *M* which is the case within a possible world does not spread to other possible worlds; on the contrary, there is a possible world where objects which have *G* do not have *M*.

On this proposal, realists would allow that there are two types of possible worlds: worlds where every object which has *G* has *M* and worlds where no object which has *G* has *M*. The type of world that would *not* be accepted is this: a world in which some objects that have *G* have *M* and some objects that have *G* do not have *M*, what Blackburn calls 'mixed worlds'. In the light of the discussion in the last section, it is also clear that realists should not accept the existence of mixed worlds, since it seems difficult to reconcile with the idea that an object's moral property depends on *G*.

Blackburn argues that the denial of mixed worlds confronts realists with what he calls 'the explanatory problem'. Given that there are possible worlds in which every object which has *G* has *M*, but also a possible world in which no object which has *G* has *M*, realists have to explain why this combination is not possible *within* a possible world. If the combination of, on the one hand, having *G* and *M*, and, on the other hand, having *G* and

---

<sup>41</sup> There might be a number of sets of non-moral properties for which this holds. Consequently, a number of necessary implications of this kind might hold.

not having M is allowed as regards all possible worlds taken together, there does not seem to be any good reason why this combination is not allowed to hold also within the same possible world. In Blackburn's words, realists have difficulties explaining 'the ban on mixed worlds'.

It is weak supervenience together with (**P**) that generates the explanatory problem, and the reason seems to be that this conception of dependence is a kind of hybrid. If an object's moral property M depends on G, then objects which have G should have M everywhere *or* the co-existence of G and M is not a matter of dependence, but rather of coincidence, in which case the combination of G and the presence or absence of M should not be excluded anywhere. In Blackburn's words: 'Supervenience becomes, for the realist, an opaque, isolated, logical fact for which no explanation can be proffered.'<sup>42</sup>

As far as I see, there is no reason for realists to be perplexed by the explanatory problem as such, since it is more of a dramatisation of the observation that weak supervenience is insufficient to guarantee dependence than a problem of its own. The reason why realists have the explanatory problem if they adopt weak supervenience together with (**P**) is that they then are not in the position to claim that the correlation between a set of non-moral properties G and a moral property M holds across all possible worlds. As a consequence, they are not in the position to make plausible that M depends on G. To point out that realists have difficulty explaining the ban on mixed worlds if they adopt weak supervenience and (**P**) is thus, as I see it, merely a way to illustrate that it is unreasonable to adopt weak supervenience if one wants to maintain that a moral property depends on a set of non-moral properties. This view finds support if we consider strong supervenience. On the assumption that a moral property M strongly supervenes on a set of non-moral properties G, it follows that there is no possible world where an object has G but does not have M. Consequently, if realists adopt strong supervenience, they do not face the explanatory problem. They may then also be in the position to claim that M depends on G.

Nevertheless, according to Blackburn, realists face a dilemma when it comes to supervenience: if realists adopt strong supervenience this leads to reduction, but if they adopt weak supervenience together with (**P**), the

---

<sup>42</sup> Blackburn (1993 (1971)), p. 119.

explanatory problem arises.

How should realists respond to the supposed dilemma? Above I have argued that realists should agree with Blackburn about step (i); realists should prefer strong to weak supervenience. Realists should also accept step (iii); as we have seen, if realists embrace the combination of weak supervenience and (P), the explanatory problem seems difficult to avoid. But let us ask whether realists should accept step (ii): does strong supervenience result in reduction?

The reason why Blackburn believes that strong supervenience results in reduction is that it together with the assumption that there is an object which has a moral property M entails that an implication like (N) is true. Moreover, he indicates that he believes that implications of this kind are false. I think both these contentions should be denied.

That an implication like (N) is true does not result in reduction, at least not in the sense this notion is understood here. The most important reason is that the truth of such an implication does not mean that a necessary biconditional involving a moral property M and a non-moral property holds. However, that such a biconditional holds is a necessary condition for property identity and hence for reduction.

Nevertheless, even if the truth of an implication like (N) is compatible with a moral property being irreducible, it might be thought that there is a persuasive argument to the effect that there are no true implications of that kind. There seem to be two main candidates for such an argument: Hume's law or Moore's open question argument.

Hume's law is often formulated in the dictum 'Is does not imply Ought'. The truth of an implication like (N) is incompatible with the spirit of Hume's law, since it means that an implication from a set of non-moral properties to a moral property holds. Hume's law seems to be the summing-up of an argument the main premise of which is that moral judgements do not consist in cognitive states, such as beliefs, but rather in non-cognitive and motivational ones, such as desires.<sup>43</sup> However, this premise is quite controversial and realists should dispute it. In fact, in chapter 7 I argued that

---

<sup>43</sup> Hume (1978 (1888)), pp. 468–470. Cf. Mackie (1980), pp. 61–63. According to another interpretation, Hume reminds us that a moral sentence cannot be inferred from a non-moral sentence without a premise which connects these two types of sentences. However, this view does not support the view that no implication like (N) holds; rather, such an implication could work as such a premiss.

the main argument for it—internalism—is mistaken.

The open question argument may seem to provide a stronger support to Blackburn's view that an implication like **(N)** does not hold. In chapter 2, I argued that an amended version of this argument is plausible. However, it is important to note that the open question argument concerns *analytically* necessary truths. Accordingly, in order for it to have any consequence for the truth of an implication like **(N)**, 'necessary' in **(N)** has to be understood as 'analytically necessary'. That is, **(N)** has to be read as **(N<sub>a</sub>)**. (The corresponding version of **(P)** would be **(P<sub>a</sub>)**.) This means that if 'necessary' in **(N)** is understood in some other way, e.g. as 'synthetically necessary', Blackburn cannot rely on the open question argument to show that no implication like **(N)** holds and, as a consequence, that realists should not adopt strong supervenience. It should also be mentioned that Blackburn indicates that it is if **(N)** is read as **(N<sub>a</sub>)** that he finds it most problematic.<sup>44</sup>

However, Blackburn believes that realists cannot read **(N)** in a non-analytic way. His argument for this contention takes its point of departure in the view that it is an analytic truth that the moral claims depend on natural claims. In a 'modern idiom', Blackburn says, to deny an analytically true proposition would be 'constitutive of lack of competence with the vocabulary' in question.<sup>45</sup> Supervenience is connected to competence in this way: 'It seems to be a conceptual matter that moral claims supervene on natural ones. Anyone failing to realize this, or to obey the constraint, would indeed lack something constitutive of competence in the moral practice.'<sup>46</sup>

Blackburn believes that realists should agree with his point about competence in respect of moral discourse and accordingly read the first occurrence of 'necessary' in a supervenience claim as 'analytically necessary'. He then argues that if realists read the occurrence of 'necessary' in weak supervenience as 'analytically necessary', this would not make any significant difference for the realist position vis-à-vis supervenience. He says that realists have to avoid claiming that an implication like **(N<sub>a</sub>)** holds, and have, by contrast, to maintain that a claim like **(P<sub>a</sub>)** does hold. Thus, they would face the explanatory problem again, but now on an analytic level. Blackburn then continues by arguing that strong supervenience should not either be adopted by realists because they would then be committed to the truth of an

---

<sup>44</sup> Blackburn (1993 (1985)), p. 137.

<sup>45</sup> Blackburn (1993 (1985)), p. 136.

<sup>46</sup> Blackburn (1993 (1985)), p. 137.

implication like  $(N_a)$ . I have difficulty following Blackburn's reasoning on this point, and it certainly deserves further discussion.<sup>47</sup> However, Blackburn can be understood to presume that if realists are convinced by his point in relation to competence with respect to moral discourse and thus read the *first* occurrence of 'necessary' in strong supervenience as 'analytically necessary', then they have to read the *second* occurrence of 'necessary' in the same way. This would mean that realists are committed to the truth of an implication like  $(N_a)$ , and no such implication holds according to the reasoning earlier.

We should however question Blackburn's presumption. That the first occurrence of 'necessary' in the strong supervenience is read as 'analytically necessary' does not imply that the second occurrence of 'necessary' has to be read in the same way. This means that the following reading of strong supervenience is open to realists:

It is analytically necessary that, for any object  $x$ , and for any moral property  $M$ , if  $x$  has  $M$ , then there is a set of non-moral properties  $G$  such that **(a)**  $x$  has  $G$ , and **(b)** it is synthetically necessary that, for any object  $y$ , if  $y$  has  $G$ , then  $y$  has  $M$ .

Needless to say, this is the realist formula without the condition **(ii)**. It seems reasonable to assume that this alternative should be open to realists because they may want to claim that it is analytically necessary that if an object has a moral property, it has some set of non-moral properties such that it necessary that if an object has that set, it has the moral property, but at the same time deny that the latter holds with *analytic* necessity.<sup>48</sup> Indeed, I have argued above that this is what realists should say. As the first occurrence of 'necessary' in this claim represents analytic necessity, it is part of being linguistically competent in respect of moral terms to know that these terms are to be applied in such a way that it is recognised that moral properties depend on non-moral properties. Thus, realists can claim with Blackburn that anybody not acknowledging this lacks competence with

---

<sup>47</sup> I think Nick Zangwill comes closest to an adequate interpretation of Blackburn's argument; see Zangwill (1995), pp. 240–262.

<sup>48</sup> James Dreier has proposed a similar claim in response to Blackburn's argument: Dreier (1992), pp. 20–21. However, he wants to combine it with the idea that moral properties are identical with natural properties. Cf. Danielsson (2001), p. 95, and McLaughlin (1995), p. 27.



regard to moral vocabulary.<sup>49</sup> But although it is part of this competence to know that moral terms are to be applied in such a way that it is acknowledged *that* moral properties depend on non-moral properties, it need not be part of this competence to know *which* non-moral properties the moral properties depend on. This is reflected in the fact that the second occurrence of ‘necessary’ does not represent analytic necessity. That is, there is no analytically necessary implication from a particular set of non-moral properties to a moral property. This version of the strong supervenience does thus not commit realists to the truth of an analytically necessary implication like (N<sub>a</sub>). If realists adopt this version of strong supervenience, they can consequently evade the argument related above.

To sum up, realists can avoid the dilemma put forward by Blackburn and maintain strong supervenience.

#### 8. *Kim: Strong or Weak Supervenience—another Dilemma*

In the preceding section, we saw that Blackburn argues that strong and weak supervenience confront realists with a dilemma. The dilemma is, roughly put, this: if realists adopt weak supervenience, they face the explanatory problem, whereas if they adopt strong supervenience, they accept a claim that leads to reduction. I also mentioned that the explanatory problem is not an independent problem, but rather a dramatisation of a point made in section 7, namely that weak supervenience is too weak to warrant the claim that an object’s moral property depends on a set of its non-moral properties. Thus understood, the dilemma put forward by Blackburn is structurally analogous to a dilemma suggested by Jaegwon Kim. That dilemma is quite general and applies not only to moral and non-moral properties, but to any attempt to state a non-reducible dependence relation between properties in terms of supervenience. The dilemma is, roughly put, the following: if weak supervenience is adopted, dependence between properties cannot be accounted for, since this kind of supervenience is too weak, whereas if strong supervenience is adopted, an important step towards reduction has been taken. Although the dilemmas proposed by Kim’s and Blackburn’s exhibit essentially the same structure, Kim’s argument for the second horn is, as we shall see, different, and, I

---

<sup>49</sup> Here I ignore that Blackburn, unlike me, believes that is ‘a conceptual matter that moral claims supervene on *natural* ones’ (Blackburn (1993 (1985)), p. 174; italics added), as this difference has no significance for the present point.

think, stronger.

The reason Kim gives for the first horn was considered in section 7, and I will not comment on it further. The reason Kim gives for the second horn is, in brief, that on the assumption that disjunction is a proper way of forming properties—that is, if a disjunction of properties itself is a property—there is a disjunctive property involving all the different base properties for a given strongly supervenient property. It follows that there is a necessary biconditional involving the supervenient property and the disjunctive property.<sup>50</sup>

Applied to the relation between moral and non-moral properties and formulated in terms of the terminology adopted above, Kim's argument can be understood in the following way.<sup>51</sup> We may start by assuming that conjunction is a proper way of forming properties. This means that if a set of non-moral properties of the kind referred to in strong supervenience consists of a conjunction of non-moral properties, it is itself a non-moral property.<sup>52</sup> Thus understood, strong supervenience says the following: it is necessary that if an object has a moral property M, it has a non-moral property G which is such that the following holds: necessarily, whatever object has G has M. Let us call the disjunction of all the non-moral properties for which this holds 'D<sub>M</sub>'. Suppose we accept that disjunction is a proper way of forming properties and hence that D<sub>M</sub> is a property. That an object has this disjunctive property then necessarily implies that it has M. Thus, the following implication, the 'first part' of the biconditional, holds: necessarily, for any object x, if x has D<sub>M</sub>, x has M. Now, recall again that, according to strong supervenience, it is necessary that if an object has M, it has a non-moral property G such that, necessarily, whatever object has G has M. That is, it is necessary that if an object has M, it has some non-moral property for which this holds. Such a non-moral property is one of the disjuncts in D<sub>M</sub>. Suppose again that we accept disjunctive properties. We

---

<sup>50</sup> See e.g. Kim (1993 (1984)), pp. 70–78, and Kim (1993 (1990)), pp. 150–155.

<sup>51</sup> As already mentioned, Kim's argument is quite general and would presumably affect all areas where there is an interest in claiming that there is a non-reducible dependence relation between properties. It is also widely discussed, particularly in the philosophy of mind. The bearing of the argument on moral properties is also considered by a number of authors; see e.g. Depaul (1987), p. 434; Jackson (1998), pp. 122–123, Shafer-Landau (2003), pp. 93–94, and Wedgewood (1999), pp. 209–210.

<sup>52</sup> That properties can be closed under conjunction is presupposed in Kim's argument. In the present context, I will for the sake of the argument take it for granted that this assumption is correct.

could then state an implication which is the ‘second part’ of the biconditional: necessarily, for any object  $x$ , if  $x$  has  $M$ ,  $x$  has  $D_M$ . The conjunction of the two implications gives us the following necessary biconditional:

It is necessary that, for any object  $x$ ,  $x$  has  $M$  if and only if  $x$  has  $D_M$ . Kim’s argument thus seems to show that strong supervenience entails a necessary biconditional involving a moral property  $M$  and a non-moral property  $D_M$ . If a biconditional of this type is true, a necessary condition for property identity is fulfilled. It might then be thought that with strong supervenience an important step towards reduction has been taken. In particular, it might be suspected that  $M$  is identical with  $D_M$ . By contrast, weak supervenience does not entail a necessary biconditional of this sort since it lacks the required modal force. But as we saw in the last two sections, weak supervenience is not an option for realism. It should be mentioned that Kim does not claim that a necessary biconditional of the mentioned kind implies that a property is identical with a disjunctive property or that the former has been reduced to the latter. Rather, he believes that it implies ‘the possibility of reducing the supervenient to the subvenient’.<sup>53</sup> It is not clear to me what further conditions Kim believes should be satisfied for property identity or reduction to be the case.

Kim’s argument raises a number of fundamental metaphysical problems that I cannot deal with here. However, I would like to respond to it in the following way.

First it is important to keep in mind that Kim’s argument rests on a controversial presumption: that properties are closed under disjunction so that a disjunction of properties, such as  $D_M$ , itself is a property. As a matter of fact, as is pointed out in a recent paper, the prevalent view in the literature seems to be that this is not the case.<sup>54</sup>

Moreover, as D. M. Armstrong in particular points out, there are some intuitively striking arguments against the existence of disjunctive properties.<sup>55</sup> Armstrong calls attention to the fact that the notion of disjunctive properties offends against the principle that a property is ‘identical in its different particulars’.<sup>56</sup> Suppose an object has a certain

---

<sup>53</sup> Kim (1993 (1990)), p. 153.

<sup>54</sup> Clapp (2001), p. 112.

<sup>55</sup> See e.g. Armstrong (1978), pp. 19–23.

<sup>56</sup> Armstrong (1978), p. 20.

property A, but lacks another property B, and that another object has B, but lacks A. It then seems paradoxical to claim that this means that they have the same property. Yet, they would if properties can be closed under disjunction, since they would have the disjunctive property ‘A or B’. Armstrong thinks it would be ‘laughable’ to maintain that these objects are identical in some respect in virtue of having such a ‘property’. Moreover, Armstrong points out that the notion of disjunctive properties breaks the natural link between properties of objects and the causal influence they exercise. Suppose again that an object has A but lacks B. Some of the causal influence that the object exercises might then be explainable in terms of A. The object’s causal influence is not explainable in terms of B for the simple reason that it lacks B. The object having ‘A or B’ does thus not add anything to the causal influence it has. But then it seems strange to claim that ‘A or B’ is a property. Moreover, the object that has A but not B might share no causal influence with an object that has B but not A. But then it seems strange to say that they have the ‘property’ ‘A or B’ in common.

To my mind, Armstrong’s arguments provide strong intuitive support against the existence of disjunctive properties. There are also other arguments to this effect.<sup>57</sup> However, it should immediately be admitted that some authors believe that there are disjunctive properties. What is important to note, however, is that they do not generally seem to accept that *all* disjunctions of properties are properties. (It can be mentioned that Kim himself in later a work maintains that not all disjunctions of properties are properties.<sup>58</sup>) On the contrary, it is often argued that only disjunctions of properties that fulfil a certain particular criterion have that status.<sup>59</sup> These authors disagree, however, what this criterion amounts to. Consequently, even if *some* disjunctions of properties are properties, it is an open question what criterion is correct and whether  $D_M$  fulfils it. Unfortunately, this issue raises deep metaphysical questions that I cannot discuss here.

However, suppose that a disjunction of properties such as  $D_M$  *does*

---

<sup>57</sup> According to an influential line of thought, one reason to be suspicious of the existence of disjunctive properties is that they cannot figure in the appropriate kind of laws. This argument is particularly associated with Jerry Fodor. For related arguments, see e.g. Pereboom and Kornblith (1991), pp. 125–132; Owens (1989), pp. 197–202, and Seager (1991), pp. 93–98. For other arguments against the existence of (some types of) disjunctive properties, see Teller (1983), pp. 57–61, and Zangwill (1998), pp. 151–164.

<sup>58</sup> Kim (1993 (1992)), pp. 319–322.

<sup>59</sup> See e.g. Clapp (2001), pp. 123–132; Penczek (1997), pp. 203–219, and Van Cleve (1990), pp. 230–232.

qualify as a property. This means that the necessary biconditional involving a moral property  $M$  and a disjunctive property  $D_M$  mentioned above holds. Does this mean that  $M$  is identical with  $D_M$ ? According to one view of property identity it does, because a property  $A$  being necessarily co-extensive with a property  $B$  provides both a necessary and sufficient condition for  $A$  being identical with  $B$ . However, according to the view of property identity that appears to be predominant, it does not. On this view, a property  $A$  and a property  $B$  may be necessary co-extensive and yet not identical.<sup>60</sup> For example, the property of being a closed straight-sided figure having three sides (being trilateral) and the property of being a closed straight-sided figure having three angles (being triangular) are necessary co-extensive. However, as for example Elliott Sober argues, being trilateral and being triangular do not seem to be the same property.<sup>61</sup> Thus, it seems reasonable to maintain that even if necessary co-extensiveness provides a necessary condition for property identity, it does not provide a sufficient condition. Hence, the biconditional above might hold without  $M$  and  $D_M$  being identical. Unfortunately, what constitutes necessary and sufficient conditions for property identity raises yet another of the fundamental metaphysical questions that I am not able to discuss here.<sup>62</sup> This means that whether Kim's argument provides evidence that a moral property  $M$  is identical with a disjunctive non-moral property  $D_M$  cannot be settled in the present work. However, as we have seen, there are a number of arguments to the effect that this is not the case. There are thus reasons to believe that strong supervenience is compatible with realism and that realists consequently can maintain this version of supervenience.

However, suppose that Kim's argument succeeds in demonstrating that a moral property  $M$  and a disjunctive non-moral property  $D_M$  are identical. This would mean that realism is false since  $M$  would not be

---

<sup>60</sup> See e.g. Achinstein (1974), pp. 266–267; Marras (1993), pp. 286–287; Shafer-Landau (2003), pp. 90–91; Smith (1999), pp. 107–108, and Sober (1982b), pp. 183–189.

<sup>61</sup> Sober (1982b), pp. 185–186. See also e.g. Shafer-Landau (2003), p. 91. For an opposing view, see Jackson (1998), pp. 125–127. For other examples of necessarily co-extensive non-identical properties, see e.g. Achinstein (1974), pp. 266–270.

<sup>62</sup> It might be proposed that a further necessary condition for property identity is sameness of causal powers; see e.g. Achinstein (1974), pp. 266–272. However, it is controversial whether sameness of causal powers is a necessary condition for property identity. Moreover, even if it is accepted that it is, it is controversial whether necessary co-extensiveness and sameness of causal powers constitute necessary and jointly sufficient conditions for properties to be identical.

distinct from a non-moral property.<sup>63</sup> I have two comments on this result. First, I think we still would have the intuition, indicated by Armstrong's arguments, that M is distinct from  $D_M$ . The argument that would have us conclude that there are disjunctive properties, and so that M is identical with  $D_M$ , would presumably be felt as a rather abstract philosophical argument which has little to do with how we ordinarily conceive of these matters. Second, if the mentioned situation occurs, we should distinguish between two kinds of reductionism with different strengths, where the weaker one may draw on the points made in support of realism in earlier chapters. According to *strong reductionism*, corresponding to what is called 'reductionism' above, a moral property M is identical with a certain non-disjunctive non-moral property. According to *weak reductionism*, a moral property M is identical with a disjunctive non-moral property  $D_M$ . However, weak reductionism retains a certain asymmetry between M and the disjuncts contained in  $D_M$ . Especially, if an object has one of the disjuncts in  $D_M$ , it has M, but the converse does not follow; what follows is merely that if an object has M, it has *some* disjunct in  $D_M$ , not that it has any particular of these disjuncts. Accordingly, advocates of this view might appeal to this asymmetry and argue that weak reductionism has some of the explanatory advantages suggested for realism in previous chapters. However, I will not examine to what extent weak reductionism can exploit what was said there.

To sum up, there are reasons to believe that realism can avoid the dilemma proposed by Kim and maintain strong supervenience. However, whether realism ultimately succeeds in doing so depends on fundamental metaphysical issues that have not been discussed here.

### 9. Mackie's Arguments from Queerness

Among the most discussed arguments against the existence of moral properties are J. L. Mackie's arguments from queerness. One of these arguments concerns the dependence of moral properties on non-moral properties and is thus directly related to the main topic of this chapter. In this section, I will discuss this argument, but I will also take the opportunity of commenting on those of Mackie's queerness arguments that are not

---

<sup>63</sup> Moreover, if this is correct, the realist formula would be self-contradictory, since it, in virtue of (i), would imply that a moral property is identical with a non-moral property and, in virtue of (ii) would deny that such an identity holds.

immediately concerned with the dependence relation.

Mackie takes the queerness arguments to provide support to his error-theory. Before he presents his queerness arguments, he argues that there are moral sentences which involve moral terms in virtue of which moral properties are ascribed to objects. If the queerness arguments are correct, there are no moral properties and all moral sentences of this kind are false; hence, error-theory is established. However, on the view that moral sentences do not ascribe moral properties to objects, these arguments may instead be taken to support non-cognitivism. Mackie's queerness arguments are in other words not necessarily connected to error-theory, but can be appealed to by other meta-ethical positions which deny the existence of moral properties.

Mackie considers there to be two arguments from queerness: one metaphysical and one epistemological. However, at closer inspection Mackie presents at least three metaphysical arguments and one epistemological argument. Furthermore, one of the metaphysical arguments has at least two different interpretations. I will discuss these arguments in turn.

According to Mackie, the arguments from metaphysical queerness show that '[i]f there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe'.<sup>64</sup> Hence, the arguments are meant to show that if there were any objective values—of which moral properties would be a subclass—they would be metaphysically queer. As we will see above, according to some interpretations of Mackie's arguments, it is not moral properties themselves that would be queer in this respect, but rather some feature related to them, such as their depending on natural properties. However, Mackie presumes that the metaphysical queerness pertaining to moral properties provides reason not to allow such properties in ontology. Hence, in Mackie's view it should make us deny the existence of moral properties.

One way to defend the existence of moral properties against Mackie's contention would be to question the mentioned presumption and argue that the existence of moral properties can be maintained even if the metaphysical queerness arguments are correct, that is, even in case moral properties were

---

<sup>64</sup> Mackie (1977), p. 38.

metaphysically queer.<sup>65</sup> However, I believe that realists can argue that the metaphysical queerness arguments fail and that Mackie therefore has not provided reason to deny the existence of such properties.

*The first argument from metaphysical queerness.* According to this argument, moral properties would be metaphysically queer because they would be, as Mackie puts it, ‘authoritatively prescriptive’, they would be ‘intrinsically action-guiding and motivating’ and have ‘to-be-doneness and not-to-be-doneness somewhat built into’ them.<sup>66</sup> In the present context, this feature seems most relevant in relation to rightness and wrongness. If I understand Mackie correctly, the authoritative prescriptivity of rightness would mean that if a person knows that an action is right, it follows that she is motivated to perform the action.

This argument does not pose any difficulty for realism, at least not for the version of this view I find most plausible. In chapter 7, I argued that realists should deny internalism, the view which says, in brief, that if a person judges that an action is right, it follows that she is motivated to perform the action. On the assumption that there are moral properties, it is reasonable to believe that such a judgement consists in a belief. Given this assumption, it is reasonable to believe that Mackie’s notion of authoritative prescriptivity implies internalism. The reason is that if a person *knows* that an action is right, she *believes* that it is right. The arguments directed against internalism in chapter 7 can then be directed against the notion of authoritative prescriptivity.<sup>67</sup>

Moreover, it can be argued that even if realists adopt internalism, they would not be vulnerable to Mackie’s argument. As I understand Mackie’s notion of authoritative prescriptivity, it says that if a person knows that an action is right to perform, and hence the judgement is *true*, it follows that she is motivated to perform the action. Internalism says that if a person judges that an action is right, it follows that she is motivated to perform the action, even if her judgement is not true. Internalism seems more reasonable than authoritative prescriptivity because it is difficult to see any reason why a person’s moral judgment would entail that she is motivated only in case

---

<sup>65</sup> Cf. Brink (1989), pp. 173–174. Cf. Wreen (1985), pp. 154–155. A similar observation holds for the argument from epistemic queerness.

<sup>66</sup> Mackie (1977), pp. 49, 40.

<sup>67</sup> Cf. Brink (1989), p. 172, and Wreen (1985), p. 153.



the judgement is true.<sup>68</sup> But if a person's moral judgement does not need to be true in order to be connected in the indicated way to motivation, there is no reason to believe that motivation pertains to metaphysical matters at all. In particular, there is no reason to believe that rightness, or other moral properties, would have to be 'intrinsically motivating' and have 'to-be-doneness and not-to-be-doneness built into' them.<sup>69</sup> On internalism, there is thus no reason to make reference to any magic force of moral properties to explain moral motivation. It is rather to be explained in terms of the content of moral judgements. In that case, Mackie's claim that moral properties are ontologically queer appears to be misguided.

*The second argument from metaphysical queerness.* According to this argument, moral properties would be metaphysically queer because they would be *non-natural properties*.<sup>70</sup>

Unfortunately, Mackie does not make clear how he conceives of the distinction between natural and non-natural properties. However, the quotation above, together with other things Mackie says, indicates that he thinks that non-natural properties are different from natural properties in that they are not part of nature or 'the universe'. He claims, moreover, that moral properties would have to be something like Plato's forms. I think one reasonable way to understand Mackie is that he conceives of the difference between natural and non-natural properties in terms of the causal relations holding in nature. Natural properties would then be properties that are part of nature in the sense that they are involved in these causal relations. Non-natural properties, on the other hand, would be properties that are not part of nature in that sense; they are not involved in its causal relations and stand accordingly outside the causal order of nature.<sup>71</sup> Of course, this conception of the distinction between natural and non-natural properties requires further examination; however, I will not pause to provide such an account here.

In reply to Mackie's second metaphysical queerness argument,

---

<sup>68</sup> Cf. Dancy (1993), p. 1.

<sup>69</sup> Suppose my interpretation of Mackie is mistaken; more precisely, suppose he holds that a judgment to the effect that an action is right entails motivation even if the judgement is not true. However, in that case there is no reason to believe that motivation is related to metaphysical matters and that moral properties are metaphysically queer.

<sup>70</sup> Mackie (1977), p. 38.

<sup>71</sup> Cf. Baldwin (1985), pp. 34–35; Smith (1999), pp. 93–101, and Sturgeon (2003), p. 538.

understood in the indicated way, it should first be recalled that in chapter 1 I defended a conception of realism according to which moral properties might be natural properties. I also mentioned that realists may claim that moral properties are natural in at least three different senses, one of them being that moral properties are part of nature in the sense of being involved in its causal relations. Moreover, in the next chapter I will defend the view that moral properties, as realism conceives of them, are involved in causal explanations of various facts. If that line of reasoning is successful, realism would not only be compatible with the view that moral properties are natural properties in the indicated sense; this claim would be part of the most reasonable version of realism.

*The third argument from metaphysical queerness.* In contrast to the two metaphysical queerness arguments considered so far, the third one does not directly concern moral properties. It concerns rather the dependence of moral properties on natural properties. Mackie writes:

Another way of bringing out this queerness [i.e. the metaphysical queerness] is to ask, about anything that is supposed to have some objective moral quality, how this is linked with its natural features. What is the connection between the natural fact that an action is a piece of deliberate cruelty—say, causing pain just for fun—and the moral fact that it is wrong? It cannot be an entailment, a logical or semantic necessity. Yet it is not merely that the two features occur together. The wrongness must somehow be ‘consequential’ or ‘supervenient’; it is wrong because it is a piece of deliberate cruelty. But just what *in the world* is signified by this ‘because’?<sup>72</sup>

I think there are at least two different interpretations of this argument.<sup>73</sup>

According to one interpretation, Mackie argues that the dependence of moral properties on natural properties would be metaphysically queer because it is unexplainable. Mackie urges in the quoted passage that it is explained what is signified by saying that something has a moral property

---

<sup>72</sup> Mackie (1977), p. 41.

<sup>73</sup> According to a third interpretation of this argument, Mackie argues that the dependence relation between moral and non-moral properties would be queer because it would hold between properties belonging to two widely different categories: non-natural and natural properties. What Mackie suggests on this interpretation is that a relation that is supposed to hold between moral properties *qua* non-natural properties and natural properties would be different from other kinds of relations ‘*in the world*’ and hence metaphysically queer. As a consequence, we should deny that there is such a relation. And since no such relation occurs, there are no moral properties. Interpreted in this way, the third metaphysical queerness argument does not pose any difficulty for realism. As just mentioned, in the next chapter I will argue that moral properties are natural properties. In that chapter, it will also be implied that moral properties depend on natural properties.

‘because’ it has certain natural properties.<sup>74</sup> He seems to presuppose that no such account can be proffered. On this interpretation, Mackie maintains that in being unexplainable, this dependence relation would be different from any other relation we encounter between properties ‘*in the world*’ and hence metaphysically queer. Consequently, its existence should be denied. And since no such relation occurs, there are no moral properties.

Interpreted in this way, I think the third metaphysical queerness argument is unfounded, the reason being that Mackie is mistaken in presuming that the relation between moral properties and what they depend on is unexplainable. In particular, it may be argued that realists are able to provide such an account in terms of the realist formula. (Mackie assumes that moral properties would depend on natural properties, whereas on the realist formula they depend on non-moral properties. As this difference is insignificant to the present argument, I will ignore it.) However, it might rightly be argued that this rejoinder to Mackie’s argument is insufficient. On this contention, to properly explain the dependence of moral properties on non-moral properties, it is not enough to come up with a claim that describes the connection between these properties. The dependence relation also has to be explained in the sense that evidence is presented which provide reasons to believe that the proposed dependence claim actually holds. However, realists are able to explain the dependence relation also in this sense. In chapter 4, I argued that the realist formula finds support in our considerations regarding the questions posed in chapters 2 and 3. Moreover, in chapters 5, 6 and 7, I argued that the realist formula is able to account for various meta-ethical issues, and this provides further support to this notion. However, it might be objected that also this rejoinder to Mackie’s argument is insufficient. On this contention, it is not sufficient that the general relation between moral and non-moral properties is explained, e.g. in the form of the realist formula. It is also required that the relation between a particular set of non-moral properties and the moral property in question is accounted for. Formulated in terms of the realist formula, it has to be explained why a particular set of non-moral properties—rather than some other one—makes an object have a certain moral property. However, as far

---

<sup>74</sup> Horgan and Timmons interpret Mackie in a similar way: Horgan and Timmons (1992a), pp. 227–230, and Horgan (1993), pp. 560–563. Their arguments raise fundamental issues that cannot be dealt with here. See Zangwill (1997), pp. 511–516, for a critique of Horgan and Timmons.

as I see, this is not a task of meta-ethics. Rather, it is something that should be settled by the correct normative theory or, put more generally, by the correct normative conception.<sup>75</sup>

According to another interpretation of the third metaphysical queerness argument, Mackie maintains that the difficulty with the dependence of moral properties on natural properties concerns the kind of modality it would involve. The modality in this dependence relation, Mackie maintains, cannot be a matter of logical or semantic necessity; neither, he thinks, can it be a matter of mere coincidence. What Mackie claims according to this interpretation is that the modality involved in this dependence relation would be different from any other modality '*in the world*' and hence metaphysically queer. Consequently, the existence of a dependence relation that involves it should be denied. And since no such relation occurs, there are no moral properties.

As we have seen, realists may argue that moral properties depend on non-moral properties in the way stated in the realist formula. One part of the formula is **(B)**: a synthetically necessary implication from a set of non-moral properties to a moral property. Although Mackie is not explicit as to what a problematic kind of modality would be, synthetic necessity has been considered with suspicion. It might therefore be thought that since the realist formula involves this notion, it is vulnerable to Mackie's argument.

Understood in this way, the argument raises fundamental questions about modality that, needless to say, I am unable to say anything about here. However, I would like to respond to this argument with the following comments.

It should first be recalled that also synthetic reductionism employs synthetic necessity; so if this notion poses a difficulty for realism, it presumably does so for synthetic reductionism as well.

Moreover, and more importantly, it should be recalled that the notion of synthetic necessity is utilised in Kripke's and Putnam's account of the reference of natural kind terms. As far as I understand, it is quite generally acknowledged that they provide a plausible view of the reference of such terms. Furthermore, to my knowledge it is not regarded as a serious problem for this account that it employs synthetic necessity. This indicates that the fact that this notion is involved in the realist formula does not pose

---

<sup>75</sup> Cf. Brink (1989), p. 175, and Shafer-Landau (2003), pp. 95–96.

a difficulty for realism. In particular, the fact that it is employed in the mentioned account of the reference of natural kind terms suggests that the realist formula does not involve a kind of modality which is different from any other modality '*in the world*'. Hence, the realist formula does not involve a metaphysically queer kind of modality. It should further be recalled that the reason why the realist formula involves synthetic necessity is due to considerations in relation to thought experiments similar to those Kripke and Putnam use to establish their view. This supports the view that, just as the modality involved in Kripke's and Putnam's view of natural kind term, the modality involved in the realist formula is not metaphysically queer. Consequently, the fact that synthetic necessity figures in the realist formula should not make us conclude that the dependence relation described in the formula does not occur and that moral properties therefore do not exist.

It should further be mentioned that some authors, among others Putnam himself, argue that the mentioned account of the reference of natural kind terms is appropriate to other terms as well.<sup>76</sup> In fact, he believes that his view of reference 'apply to the great majority of all nouns, and to other parts of speech as well'.<sup>77</sup> Perhaps we are not prepared to agree that this view of reference is so generally applicable. However, if it applies at least to some other terms than natural kind terms, synthetic necessity may figure in accounts of the reference of terms that are more akin to moral terms than natural kind terms are. Moreover, it may then be hypothesised that synthetic necessity figures in accounts of terms which refer to properties that are irreducible. In that case, synthetic necessity might be employed in much the same way in these accounts as it is employed in the realist formula.<sup>78</sup> Mental terms might be an example. According to one view, irreducible mental properties strongly supervene on physical properties and the second occurrence of 'necessary' in this supervenience claim should be understood as synthetic necessity.<sup>79</sup> Such a claim comes quite close to the realist formula.

There are also other considerations that might lessen the worry that the modality involved in the realist formula is metaphysically queer. One

---

<sup>76</sup> See e.g. Putnam (1975), pp. 242–245. See also e.g. Copp (2000), pp. 114–116. Of course, particularly Kripke argues that a similar account is available for names.

<sup>77</sup> Putnam (1975b), p. 242.

<sup>78</sup> Cf. Brink (1989), pp. 175–176.

<sup>79</sup> See e.g. Macdonald (1989), p. 194. Cf. Kim (1993 (1984)), p. 66.

reason to be suspicious of synthetic necessity is the view that it makes up a separate kind of necessity which is quite different from that of analytic necessity. However, according to an influential view, there is only one kind of necessity of the relevant sort.<sup>80</sup> On this view, the distinction between analytic and synthetic necessity is not a distinction between two separate kinds of necessities; it is rather a distinction between the different epistemic relations we have to this one necessity.<sup>81</sup>

*The argument from epistemological queerness.* Also Mackie's epistemological queerness argument is directed against the existence of moral properties. According to Mackie, if we are aware of moral properties 'it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else'.<sup>82</sup> Hence, such a faculty would be epistemologically queer. Mackie considers this, in analogy with his line of reasoning in relation to the metaphysical queerness arguments, to provide a reason to deny that there is such a faculty. Consequently, we are not aware of moral properties. Assuming that we would be aware of moral properties if there were any, it follows that there are no moral properties.

As I mentioned in chapter 1, one of the main limitations with the present thesis is that I do not discuss epistemological issues in the way that would be required to provide a fully satisfactory defence of realism. Here I will merely provide some brief reflections that give reasons to believe that realism is not vulnerable to Mackie's epistemological queerness argument.

Mackie's reason to believe that, on the assumption that we are aware of moral properties, there would have to be a special faculty of moral perception is arguably his view that moral properties would have to be non-natural properties.<sup>83</sup> Thus understood, the idea is that since moral properties would have to be non-natural properties, they cannot be known by means of our ordinary perceptual apparatus standing in causal contact with them.

---

<sup>80</sup> See e.g. Jackson (1998), pp. 67–86, and McLaughlin (1995), pp. 26–27.

<sup>81</sup> Another reason to be suspicious of synthetic necessity is the view that it cannot be explained. However, in relation to the last point, it has been argued that synthetic necessity might be explained in terms of analytic necessity or linguistic conventions. If I have understood this view correctly, it does not say that there is no synthetic necessity; it says rather that synthetic necessity is generated by analytically necessary claims in combination with certain other premises. See e.g. Jackson (1989), pp. 80–83, and Sidelle (1989), pp. 30–40.

<sup>82</sup> Mackie (1977), p. 38.

<sup>83</sup> Cf. Brink (1989), p. 180. For another interpretation, see Sturgeon (2002), pp. 194–195.

For this reason, some sort of special faculty is needed. However, as has already been noted, realists may claim that moral properties are natural properties, among other things in the sense that they are part of nature by being involved in its causal relations. In the next chapter, I will suggest that moral properties are natural properties in this sense. My arguments will moreover suggest that the non-moral properties which moral properties depend on are natural properties. On this view, it might be the case that we can come to acquire moral knowledge in virtue of coming into causal contact with properties belonging to these two kinds. This can be the case in different ways. According to a common conception of observations, an observation is a belief which is a direct causal result of a perceptual experience.<sup>84</sup> Thus, the most apparent way in which we can acquire moral knowledge is presumably that the mentioned properties cause us to have moral observations, i.e. moral beliefs which are direct causal results of sense perceptions. However, there are also other types of facts these properties can cause and which can provide us with moral knowledge. It might be argued, for example, that there are observations which do not involve reference to any moral property which can have this function. (In the next chapter, I will mention some facts of this kind.) One important way in which we may acquire moral knowledge was mentioned in chapter 4. There I suggested that realists may hypothesise that we are able to apply a moral term correctly because there are certain reference-fixing characteristics that we associate with the term. This is made possible since the non-moral properties which make objects have moral properties are causally responsible for these characteristics and so causally regulate our use of the term by means of them. Hence, as moral realism is understood here, it is not committed to the existence of a special faculty of moral perception.

It should also be noticed that Mackie's notion of a special faculty of moral perception makes him presume that the correct view of moral justification would have to be a version of intuitionism.<sup>85</sup> Intuitionism is an instance of foundationalism, a view which says, in brief, that a belief is justified only if it is foundational, i.e. self-justifying or non-inferentially

---

<sup>84</sup> See e.g. Harman (1977), p. 5; Sayre-McCord (1996), p. 174; Tännsjö (1990), p. 54, and Werner (1983), p. 653. It should be noted that on this conception of moral observation, there need not be perceptions directly of moral properties (cf. Werner (1983), p. 666). For an account that attaches a fundamental importance to moral observations in moral justification, see Tännsjö (1990), chap. 3.

<sup>85</sup> Mackie (1977), p. 38.

justified, or inferred from some foundational beliefs.<sup>86</sup> The picture Mackie suggests is that the special faculty of moral perception provides us with moral ‘intuitions’ which have the status of foundational moral beliefs.

There are at least three difficulties with Mackie’s view of moral justification. First, a number of meta-ethicists argue that intuitionists are not committed to the existence of a special faculty of moral perception.<sup>87</sup> This means that there might be a more plausible form of intuitionism than the one Mackie sketches that realists may adopt.<sup>88</sup> Second, it might be argued that there also are other forms of foundationalism than intuitionism which realists can adopt. According to one such view, moral observations—understood as moral beliefs which are direct causal results of perceptual experiences—function as foundational beliefs. However, I will not try to evaluate the plausibility of these alternatives. Third, a number of recent meta-ethicists argue that the correct model of moral justification is not foundationalism, but some version of coherentism.<sup>89</sup> Coherentism says, in brief, that a belief is justified in so far it is part of a coherent set of beliefs, where the coherence of such a set is a function of, among other things, the explanatory relations that hold between the beliefs which are part of the set. I will not dwell on how coherentism should be understood to apply to moral justification. However, in light of what was said above, it is important to notice that the view that beliefs are justified in so far as they are part of a coherent set of beliefs does not mean that perceptual experiences do not have any part to play in moral justification. Generally, according to coherentism, perceptual experiences are relevant to the coherence of a set of beliefs by giving rise to beliefs that are part of such a set. This means that for example moral observations—understood in the indicated way—can be part of a coherent set of beliefs, e.g. in virtue of explaining other beliefs in the set and thereby enhancing its coherence.<sup>90</sup> Corresponding claims hold for other beliefs that pertain to perceptual experiences. It is also worth mentioning that on coherentism a wide range of beliefs might be relevant to

---

<sup>86</sup> My characterisation of foundationalism and coherentism follows Brink’s: Brink (1989), pp. 101–104.

<sup>87</sup> See e.g. Audi (1996), pp. 108–109; Brink (1989), pp. 109–110, and Sinnott-Armstrong (1996), p. 26.

<sup>88</sup> Robert Audi is presumably the one who has done most to develop a modern form of intuitionism; see e.g. Audi (1996), pp. 101–136.

<sup>89</sup> See e.g. Boyd (1988), pp. 199–202, 206–209; Brink (1989), chap. 5; Goldman (1988), chap. 5, and Sayre-McCord (1996), pp. 137–189

<sup>90</sup> Cf. Brink (1989), pp. 136–138, and Sayre-McCord (1996), pp. 173–174.



moral justification, e.g. philosophical beliefs and other beliefs of a markedly theoretical character.

To sum up, there are reasons to believe that Mackie's various arguments from queerness do not succeed in showing that realists are committed to the existence of phenomena—either in the form of moral properties, dependence relations or epistemological faculties—that are queer in the relevant sense. Hence, there are reasons to believe that these arguments do not establish that moral properties do not exist. However, Mackie's arguments also raise important questions that I have not dealt with here.

### **10. *Concluding Remarks***

This chapter has been devoted to a discussion of how realism should conceive of the dependence of moral properties on non-moral properties. In particular, I have defended the realist formula against a number of objections. Among other things, I have responded to two dilemmas that are thought to emerge because the realist formula entails strong supervenience, and I have responded to Mackie's arguments from queerness. However, I have also pointed out various respects in which this defence of realism is incomplete. I would like to end the chapter by recalling two important reasons why this is so. First, the dilemma Jaegwon Kim presents suggests that if strong supervenience is adopted, an important step towards reducing moral properties to non-moral properties has been taken; in particular, the argument might be taken to show that moral properties are identical with non-moral properties. Although I have argued that there are reasons to believe that the adoption of strong supervenience does not have this consequence, whether this is correct depends on fundamental considerations in metaphysics that have not been dealt with here. Second, Mackie's epistemological queerness argument raises the question how moral epistemology should be understood on realism. Although I gave some reasons to believe that realism can avoid this argument, I have not provided any realist moral epistemology.

## Chapter 9

### Moral Explanations

#### 1. *Introduction*

One of the most debated issues in contemporary meta-ethics is whether moral properties have any explanatory function and particularly whether they are involved in causal explanations. According to a much debated argument put forward by Gilbert Harman, moral properties are not part of such explanations, at least not if they are irreducible, as realism claims. This contention makes him draw the conclusion that we are not justified in believing in the existence of moral properties thus conceived. The main part of this chapter is devoted to arguing that, contrary to Harman's contention, there are reasons to maintain that moral properties are involved in causal explanations, also on the assumption that they are irreducible. Consequently, realists are justified to uphold their existence. However, I will also consider the explanatory function of such properties.

In the next section, I give a brief outline of Harman's argument. As we will see, it involves as an essential premise an explanatory criterion which states a condition that an entity has to fulfil in order for us to be justified in believing in its existence. In developing his argument, Harman adduces three different criteria, which I will refer to as 'the causal criterion', 'the best explanation criterion' and 'the pragmatic criterion'. The two first criteria are the most central. In section 3, I argue that, contrary to Harman's contentions, there are reasons to maintain that moral properties fulfil the first criterion. However, Harman admits that this criterion might be too strong and proposes the second criterion. In relation to the second criterion, Harman argues that moral properties fulfil it only on condition that they are reducible. If this is correct, we would not be justified in believing in the existence of moral properties as realism conceives of them. I argue in section 4 that there are reasons to believe that moral properties fulfil the second criterion even if they are irreducible. In section 5, I maintain that Harman's argument faces a dilemma which provides reason to believe that even if irreducible moral properties do not fulfil the second criterion, this should not make us draw the conclusion that they do exist; rather, it indicates that

the criterion is too strong. In section 6, I consider the third explanatory criterion Harman appeals to, arguing that irreducible moral properties fulfil it. Having maintained that irreducible moral properties fulfil Harman's causal criteria, I discuss in section 7 how realists should conceive of the causal powers of such properties.

## 2. *An Outline of Harman's Argument*

According to Harman, there is a significant difference between scientific entities and moral properties.<sup>1</sup> In brief, while scientific entities are involved in explanations of scientific observations, moral properties are not involved in explanations of moral observations. Harman formulates this view by saying that moral properties are not 'needed' to explain moral observations or that they are 'totally irrelevant' to the explanations of them. As mentioned in the last section, Harman appeals to two main explanatory criteria. Their common essence is, roughly put, that in order for us to be justified in believing that a certain entity exists, it has to be involved in causal explanations of our observations with regard to it. The idea behind this assumption is presumably that for us to be justified in believing in the existence of a certain entity, we have to have some evidence of its existence. On a general empiricist notion, we have such evidence only if there is reason to believe that it has some kind of effect on our observations or helps to account for such an effect. And if it has such an effect, or helps to account for such an effect, it is involved in the causal explanations of our observations with regard to it. From the assumed difference between scientific entities and moral properties together with the explanatory criteria, Harman draws the conclusion that we are not justified in believing in the existence of moral properties. By contrast, we are justified in believing in the existence of scientific entities, at least as far as the criteria are concerned.

As indicated, Harman argues that we are not justified in accepting the existence of moral properties because they are totally irrelevant to the explanation of moral observations. An observation is, in Harman's terminology, a belief formed as a direct result of a perceptual experience.<sup>2</sup> However, it is clear that Harman takes the argument to establish that moral

---

<sup>1</sup> Harman presents the argument in Harman (1977), chap. 1, and Harman (1986a), pp. 57–68.

<sup>2</sup> Harman (1977), p. 5.

properties are totally irrelevant to the explanation of other facts as well.<sup>3</sup> In what follows, I will at times draw on this more general formulation of the argument.

### 3. *The Causal Criterion*

Harman introduces his argument by exemplifying the difference he thinks holds between scientific entities and moral properties regarding causal explanations of observations. Suppose a physicist who sees a vapour trail in a cloud chamber comes to believe ‘There goes a proton’. Harman claims that, given certain assumptions mainly about the physicist’s scientific beliefs, a reasonable explanation of this observation is that a proton actually moves through the cloud chamber. However, suppose a person who sees some kids pour gasoline on a cat that they then set on fire comes to believe ‘That action is wrong’. In such a case, Harman claims, a reasonable explanation of this observation is not that the action actually has the moral property of being wrong. Rather, a moral observation is to be explained by referring primarily to the person’s ‘moral sensibility’, i.e., roughly, the moral beliefs she already has.<sup>4</sup> Harman formulates the difference between the two cases in the following way:

Facts about protons can affect what you observe, since a proton passing through the cloud chamber can cause a vapour trail that reflects light to your eye in a way that, given your scientific training and psychological set, leads you to judge that what you see is a proton. But there does not seem to be any way in which the actual rightness or wrongness of a given situation can have any effect on your perceptual apparatus. In this respect, ethics seems to differ from science.<sup>5</sup>

Harman’s stress on ‘cause’ and ‘effect’ gives reason to hypothesise that he appeals to the following criterion:

*The causal criterion:* We are justified in believing that a certain entity exists only if we are justified in believing that it causes, and hence causally explains, our observations with regard to it.<sup>6</sup>

---

<sup>3</sup> Cf. Sturgeon (1984), p. 54.

<sup>4</sup> On both these accounts, we also have to refer to various additional facts in order to explain a moral observation, e.g. facts about the perceptual apparatus of the person in question. I will consider reference to these facts as implied in what follows.

<sup>5</sup> Harman (1977), pp. 7–8.

<sup>6</sup> Cf. Sayre-McCord (1988), p. 263. I will take ‘entity’ to cover e.g. properties. It might be argued that what ultimately have causal powers are not properties themselves, but their instances or exemplifications. (See e.g. Marras (1993), pp. 293–294.) As I do not think this qualification is relevant in the present context, I will ignore it.

Harman's argument would then be that since moral properties, in contrast to scientific entities, do not fulfil the causal criterion, we are not justified in accepting their existence.

A general problem with Harman's argument is that it is difficult to see that he offers any clear reasons to believe that scientific entities causally explain scientific observations, whereas moral properties do not causally explain moral observations. The example Harman gives does not, as far as I understand, prove the alleged difference between scientific entities and moral properties; rather, it illustrates what should be proved.

It is important to notice from the outset that we provide explanations in terms of moral properties. Moreover, it is important to notice that in doing so, we assume that these properties are causally responsible for various facts. Thus, it seems to be part of our conception of moral properties that they make things happen, that they have effects on how things are going. As Nicholas L. Sturgeon argues, there are several types of moral explanations where this assumption is implied.<sup>7</sup> In one quite common kind of cases, we explain a person's actions by referring to a moral property related to her character. We may for example claim that a person performs a certain action—helps people in need in spite of considerable costs to herself, say—because she is good. In another kind of cases, we explain people's responses to actions by referring to a moral property these actions are thought to have. We may for example claim that a person approves of a certain action because it is right. Another example belonging to this category occurs when a person's moral belief about an action is explained by reference to its moral property. (It is this kind of moral explanations Harman focuses on.) In a third kind of cases, we explain an event in a society by referring to a moral property of one of its institutions. We might for example claim that a revolution broke out because the legal system is unjust.

Moreover, moral explanations often seem to be confirmed, and this can be taken to support the view that moral properties actually are causally explanatory. Suppose, for example, that it is claimed that a certain person is good. Considering the assumed relation between being a good person and helping people in need despite considerable costs to oneself, it might be expected that if the person in question is given the opportunity, and some

---

<sup>7</sup> Sturgeon (1984), pp. 63–65. See also Sturgeon (1986b), pp. 122, 124–125. For an account of explanations in terms of moral rules, see Sayre-McCord (1992), pp. 55–71.

further conditions are satisfied, she will help people in need even if it would mean considerable costs to herself. Predictions of this kind often turn out to be correct. This can be taken to confirm the view that a person's goodness is causally responsible for her actions. Similar considerations hold for other moral properties and for the other kinds of moral explanations.

These considerations do not demonstrate that moral properties have any causal powers; after all, through history a number of explanations have seemed plausible which we nowadays consider to be mistaken. However, it seems fair to claim that they give us *prima facie* reasons to maintain that moral properties are causally explanatory. It is in other words a belief we are justified in maintaining in the absence of successful arguments to the contrary. Moreover, to give up such a fundamental belief requires strong reasons.

Furthermore, there is a certain inconsistency in Harman's reasoning which is relevant in the present context. Before Harman presents his case against the existence of moral properties, he makes clear that all observations are 'theory laden'. This idea is familiar from contemporary philosophy of science. It means, roughly put, that what a person observes to a certain extent is a function of her prior beliefs which are part of the 'theory' she holds about the world. This, Harman claims, is true both of scientific and moral observations. Harman believes, as we have seen, that moral observations are explained in terms of the person's prior moral beliefs; no reference to moral properties is needed. But as regards scientific observations, he believes that they are explained in terms of prior scientific beliefs *and* scientific entities. However, since Harman fails to point to any relevant difference between the scientific case and the moral case, he should say that what holds for the scientific case also holds for the moral case, and vice versa. That is, either he should say that both scientific and moral observations are explained by prior beliefs, or that both are explained also by entities in the world. Since he presumably is not prepared to draw the first conclusion, he should opt for the second one.

According to contemporary philosophy of science—influenced by the works of Pierre Duhem and W. V. O. Quine—we should assume that at least some of our theory laden prior beliefs are true. The reason is, in brief, that since we cannot test any hypothesis in isolation from what we otherwise believe, we have to presume that at least some of the beliefs that

are part of the theory we currently hold are true. If we make this assumption, we do not have to conclude that scientific observations are explained merely in terms of prior scientific beliefs without any reference to scientific entities. For example, if we assume that the physicist's scientific beliefs about, among other things, the relation between vapour trails and protons are true, there is an explanation of her observation 'There goes a proton' which involves the existence of protons. What is notable is that if we make the corresponding assumption about our moral beliefs, we reach a similar conclusion as regards moral properties. If we assume that at least some of our moral beliefs are true, we need not rest content with explanations merely in terms of moral beliefs, but are in the position to provide explanations of our moral observations which involves the existence of moral properties.<sup>8</sup> For example, if we assume the truth of our beliefs about, among other things, the wrongness of causing intense pain to animals just because it is fun, there is an explanation of the person in the example above making the observation 'That action is wrong' which involves the existence of wrongness.

It might be objected that we are justified in assuming that certain beliefs which are part of scientific theories are true, but that we are not justified in making the corresponding assumption about moral beliefs. There are mainly two problems with this response. First, to deny that we should assume the truth of certain prior moral beliefs would be to make exception to a principle that has found widespread acceptance in contemporary philosophy of science, and it is difficult to see why it should not apply to the moral area as well.<sup>9</sup> It may also be noted that Harman has not argued for this view. Second, if Harman would maintain that we should not presume the truth of certain moral beliefs in the way we assume the truth of certain scientific beliefs, it is difficult to see that he has produced an independent argument against the existence of moral properties. In order to be able to maintain the view that the moral beliefs are not true, he has to have support from some meta-ethical argument, e.g. some argument discussed in preceding chapters. But then it is difficult to see that his argument from explanation

---

<sup>8</sup> Cf. Brink (1989), pp. 185–186; Sturgeon (1984), pp. 67–72, and Werner (1983), pp. 665–673. Of course, we also have to presume other background assumptions than those I mentioned here.

<sup>9</sup> Note that we do not have to assume that all, or most, of our moral beliefs are true; we only have to assume that some of them, perhaps the most fundamental, are.

adds any reason to believe that we are not justified in accepting the existence of moral properties. After all, if we assume that our moral beliefs do not hold true, this would in itself suggest the non-existence of moral properties.<sup>10</sup>

Another difficulty with the argument under discussion is that the causal criterion seems too strong. Harman himself recognises this difficulty and exemplifies it with mathematics.<sup>11</sup> We hardly believe that the subject matter of mathematics, such as numbers, causally explain our mathematical beliefs. But from this we do not draw the conclusion that there are no numbers, that mathematical sentences are false and that we cannot have mathematical knowledge.<sup>12</sup>

To sum up, there are *prima facie* reasons to maintain that moral properties are causally explanatory, and Harman has not presented any successful argument against this view. Thus, there is reason to maintain that moral properties fulfil the causal criterion. Moreover, even if they would not, this would not show that we are not justified in believing in their existence, since the criterion is too strong.

#### 4. *The Best Explanation Criterion*

Harman believes, as we have seen, that the causal criterion might be too strong, among other things because of the implication it has for mathematics. However, he still thinks there is a relevant difference between mathematics and morality:

Since an observation is evidence for what best explains it, and since mathematics often figures in the explanations of scientific observations, there is indirect observational evidence for mathematics. There does not seem to be observational evidence, even indirectly, for basic moral principles. In explaining why certain observations have been made, we never seem to use purely moral assumptions. In this respect, then, ethics appears to differ not only from physics but also from mathematics.<sup>13</sup>

It might be hypothesised that Harman appeals to the following criterion:

---

<sup>10</sup> Cf. Sturgeon (1984), pp. 56–57.

<sup>11</sup> Harman (1977), pp. 9–10. Harman's reference to mathematics may have us revise the causal criterion. The reason is that it may seem strange to assume that numbers are entities in any reasonable sense. In that case, the truth of mathematical sentences or mathematical knowledge would not be threatened by the fact that numbers do not pass the causal criterion. To avoid this difficulty, we might formulate the causal criterion in terms of truths rather than in terms of entities.

<sup>12</sup> For other examples, see Sayre-McCord (1988), p. 266.

<sup>13</sup> Harman (1977), p. 10.



*The best explanation criterion:* We are justified in believing that a certain entity exists only if the assumption that it exists is part of ('figures in') the best explanations of our observations with regard to it.<sup>14</sup>

Assuming this criterion, the argument would be the following. Even if matters belonging to mathematics, e.g. numbers, do not cause anything themselves, mathematical assumptions are part of the best explanations of certain scientific observations. By contrast, the assumption that there are moral properties is not part of the best explanations of moral observations. According to the criterion, it follows that we are not justified in believing in their existence.

Harman believes that this argument can be met on condition that moral properties are reducible. He suggests that moral properties might be reducible to non-moral properties that are referred to in the best explanations of our moral observations.<sup>15</sup> In that case, a necessary condition for us being justified in accepting the existence of moral properties is fulfilled. However, Harman does not think the argument can be met if moral properties are conceived of as irreducible. He believes that the assumption that there are irreducible moral properties is not part of the best explanations of our moral observations; as he puts it, such properties are 'totally irrelevant' to the explanations of such observations. Consequently, irreducible moral properties do not fulfil the best explanation criterion, and we are not justified in believing in their existence.

If Harman is correct, moral properties would fulfil the best explanation criterion on reductionism but not on realism. It is therefore relevant to inquire whether Harman is correct in claiming that the assumption that there are moral properties is not part of the best explanations unless they are reducible.

The argument against the existence of irreducible moral properties issuing from the best explanation criterion has a difficulty similar to the argument issuing from the causal criterion: Harman does not offer any clear

---

<sup>14</sup> Cf. Sayre-McCord (1988), p. 267.

<sup>15</sup> Harman (1977), p. 13, and Harman (1986a), pp. 64–67. Although Harman has some doubts as to whether such a reduction is possible, this seems to be the alternative he opts for; see Harman (1986a), pp. 66–67. Harman assumes that moral properties would be reducible to *natural* properties. However, I will formulate his view in terms of non-moral properties. This is of no consequence to my arguments against Harman's views. Harman is not explicit about which kind of reduction he has in mind, but what he says suggests that he believes that a reduction of a moral property implies that it is identical with a natural property; see Harman (1986a), p. 67.

reason as to why moral properties being reducible or not should be relevant to whether their existence is assumed in the best explanations. However, in the quotation above he makes—more or less in passing—a suggestion that perhaps might be taken to bolster his view: ‘In explaining why certain observations have been made, we never seem to use purely moral assumptions.’ The argument can be understood in the following way: the fact that we do not refer to moral properties to explain moral observations shows that the assumption that there are moral properties is not part of the best explanations of these observations; at least, this holds unless they are reducible in the mentioned way. There are at least two difficulties with Harman’s contention.

Firstly, it is difficult to see that the quoted remark agrees with the facts. It seems that it has mainly two interpretations. It might mean that we never appeal to moral properties *at all* to explain moral observations, or it might mean that we never appeal *only* to moral properties to explain moral observations, but that in the cases where we refer to such properties, we always refer to non-moral properties as well. However, both these views seem to be falsified by the way in which we provide these explanations. As can be seen from the examples Sturgeon gives of moral explanations, we do refer to moral properties to explain various facts, including moral observations. And sometimes we refer only to moral properties, without mentioning any non-moral properties. (I will return to the second interpretation below.) Hence, there are *prima facie* reasons to believe that moral properties are involved in the relevant explanations irrespective of whether they are reducible or not.

Secondly, Harman’s view that the assumption that there are irreducible moral properties is not part of the best explanations does not find support in the pertinent test. As we have seen, Harman formulates his view by saying that irreducible moral properties are ‘totally irrelevant’ to explanations. Sturgeon proposes the following test of this kind of irrelevance: ‘if a particular assumption is completely irrelevant to the explanation of a certain fact, then the fact would have obtained, and we could have explained it just as well, even if the assumption had been false.’<sup>16</sup> For example, to have an indication of whether the wrongness of a certain

---

<sup>16</sup> Sturgeon (1984), p. 65. Sayre-McCord proposes a modified test: Sayre-McCord (1988), p. 272. For a test of the explanatory relevance of mental properties that resembles Sturgeon’s test, see e.g. Marras (1993), pp. 294–295.

action is completely irrelevant to the explanation of a person's belief that it is wrong, we should inquire if she would have responded in this way even if the action had not had that property. Now, as we saw in the last chapter, moral properties strongly supervene on non-moral properties. That is, it is necessary that if something has a certain moral property, it has some set of non-moral properties, such that the following holds: it is necessary that if an object has that set of non-moral properties, it has the moral property.<sup>17</sup> We might believe, for example, that if an action involves deliberately causing intense pain to an animal just for the fun of it, it is wrong. That moral properties strongly supervene on non-moral properties means that if the action had not been wrong, it would have been different as regards its non-moral properties. But, Sturgeon points out, if the action had been different in that respect, it is reasonable to assume that the person in question would not have responded as she did.<sup>18</sup> For example, if the action had not caused pain to the animal, she would not have believed that it is wrong. Thus, if the action had not been wrong, the person would not have responded as she did. Similar considerations apply to other moral properties. This line of reasoning holds irrespective of whether moral properties are irreducible or not. Thus, Sturgeon's test does not provide support to the claim that irreducible moral properties are completely irrelevant to the explanations of facts. This means, according to this test, that it has not been shown that the assumption that there are such properties is not part of the best explanations.<sup>19</sup>

---

<sup>17</sup> Instead of appealing to strong supervenience, the same point can be made in terms of the realist formula. However, as Sturgeon's point is relevant quite apart from the correctness of the realist formula, I formulate it in terms of strong supervenience. I will make use of strong supervenience also in other contexts where a general issue concerning the relation between moral and non-moral properties in relation to explanation is under discussion. However, when I consider more specifically how realists should understand the causal powers of moral properties, I will employ the realist formula.

<sup>18</sup> Of course, a person might be such that she would think that the action is wrong even if the action had quite other non-moral properties. She might for example believe that whatever young people do, it is wrong. However, this is not something Harman can presume. Cf. Sturgeon (1984), pp. 66–67.

<sup>19</sup> Sturgeon notices that it might be objected that this reasoning is question begging against Harman's argument. The reason would be that it rests on the view that the following kind of implications hold: necessarily, if something has a certain set of non-moral properties, it has a certain moral property. However, Harman might want to deny all such assumptions and claim that something may have whatever non-moral properties without having any moral property. So even if the action mentioned above causes intense pain to an animal, it is not wrong. In that case, the person in question might believe that the action is wrong although this is not the case. Then there would be

In response to objections put forward by Sturgeon, Harman has provided basically two arguments to support his claim that irreducible moral properties are completely explanatorily irrelevant.<sup>20</sup>

Firstly, Harman argues that the types of moral explanations Sturgeon refers to do not imply that the moral properties *themselves* have any explanatory function. He puts this point by saying that such an explanation does not show that it is ‘the actual *wrongness*’ of an action that is ‘generating’ a person’s moral belief that the action is wrong.<sup>21</sup> As I read Harman, he maintains that unless moral properties are reducible to non-moral properties that explain moral observations, examples such as Sturgeon’s fail to indicate that moral properties explain such observations. He also insists that in the absence of such reductions, moral observations are to be explained in terms of people’s prior moral beliefs without any reference to moral properties.

Secondly, and in close connection to the first argument, Harman argues that Sturgeon’s test does not provide an appropriate test of whether entities are completely irrelevant to explanations. Put in the terms Harman makes use of in his original formulation of the argument, his point can be stated thus: although irreducible moral properties as characterised by ‘moral epiphenomenalism’ do not qualify as completely irrelevant according to Sturgeon’s test, they are indeed so. Harman understands moral epiphenomenalism as the view that moral properties are ‘epiphenomenally supervenient on natural properties in the sense that the possession of moral properties is explained by possession of the relevant natural properties and nothing is influenced or explained by the possession of moral properties’.<sup>22</sup> What cause the fact to be explained—in our example, the person’s belief

---

reason to believe that the wrongness of the action is completely irrelevant to the explanation of her moral belief according to Sturgeon’s test. However, it seems difficult for Harman to deny the mentioned kind of assumptions. One reason is that they are a fundamental part of our moral outlook and to abandon them would require good arguments. However, Harman has not provided any arguments to this effect. Another reason is that Harman presumably is not prepared to deny the corresponding kind of assumptions in other areas, e.g. science. He might want to reply that there is a difference between morality and science that justifies us to deny such assumptions in the first case but not in the latter. This takes us back to an observation made in the last section. To make his argument plausible, Harman has to explain why we should accept prior scientific beliefs but not prior moral beliefs. Moreover, if he makes this move, it is difficult to see that he has succeeded to provide an independent argument against the existence of moral properties. See Sturgeon (1984), pp. 68–73.

<sup>20</sup> Harman responds to Sturgeon (1984), pp. 49–78 in Harman (1986a), pp. 57–68. Sturgeon responds in turn in Sturgeon (1986a), pp. 69–78.

<sup>21</sup> Harman (1986a), p. 62. Cf. Harman (1984), pp. 33–34.

<sup>22</sup> Harman (1986a), p. 63.

that an action is wrong—are according to moral epiphenomenalism the non-moral properties on which the wrongness of the action supervenes. Since the wrongness supervenes on these non-moral properties, the fact would not have occurred unless the action had been wrong. That is, if the action had not been wrong, it would not have had the non-moral properties that cause the fact, and the fact would consequently never have occurred. The wrongness would thus not qualify as completely irrelevant on Sturgeon's test. Harman stresses however that according to moral epiphenomenalism, the wrongness of the action itself does not cause any facts. As I read him, he therefore concludes that the moral property is completely irrelevant to explanations even if it does not qualify as such according to Sturgeon's test.

As far as I can see, Harman's arguments fail to provide support for his view that moral properties are explanatorily irrelevant unless they are reducible.

First I would like to comment on the dialectic of the debate and on which side that has the burden of proof. Earlier I mentioned that in providing moral explanations, we sometimes refer only to moral properties to explain various facts, without mentioning any non-moral properties. In the first argument, Harman maintains that the examples Sturgeon gives of moral explanations do not show that moral properties themselves explain moral observations. However, since we enter the debate with the mentioned assumption, it is up to those who believe that moral properties themselves are *not* involved in the relevant explanations to prove their case. As far as I can see, Harman has not provided any arguments to this effect. An analogous comment applies to Harman's second argument. He suggests that there is one view—moral epiphenomenalism—according to which irreducible moral properties themselves do not cause anything. However, the mere fact that there *is* such a position cannot establish that irreducible moral properties are not causally effective.<sup>23</sup> Moreover, Harman does not offer any argument to the effect that moral epiphenomenalism follows if moral properties are irreducible. Hence, what he says does not show that moral properties are causally ineffectual if they are irreducible.

Let us now consider more specifically the first of Harman's arguments. I think this argument reveals a conflation between the two

---

<sup>23</sup> Cf. Sturgeon (1986a), p. 75.

explanatory criteria considered so far: the causal criterion and the best explanation criterion. According to the first criterion, for us to be justified in believing in the existence of an entity, it is necessary that we are justified in believing that the entity itself is causally effective. According to the second criterion, for us to be justified in believing in the existence of an entity, it is merely necessary that the assumption that it exists is part of the best explanations. When Harman complains that the kind of examples Sturgeon's gives does not show 'the actual *wrongness*' of an action is 'generating' the belief that the action is wrong, his formulation suggest that he has the first criterion in mind. However, as he has abandoned this criterion, he cannot appeal to it.

Moreover, it is reasonable to argue that the assumption that there are moral properties can be part of the best explanations even if such properties are not causally effective themselves. Admittedly, this assumption may perhaps not be part of the best explanations in the same way in which mathematical assumptions, according to Harman, is part of such explanations. However, the assumption that there are moral properties might be part of such explanations in virtue of moral properties strongly supervening on non-moral properties that are causally effective. According to strong supervenience, it is necessary that if something has a certain moral property, it has some set of non-moral properties such that the following holds: it is necessary that if something has that set of non-moral properties, it has the moral property. Suppose that the non-moral properties on which the moral property supervenes are causally effective. This suggests that the assumption that these properties exist is part of the best explanations.<sup>24</sup> Given that the moral property strongly supervenes on these non-moral properties—so that, necessarily, if an object has these non-moral properties, it has the moral property—it is then reasonable to believe that the assumption that the moral property exists also is part of such explanations. Hence, it is reasonable to believe that moral properties can fulfil the best explanation criterion even if they are not causally effective themselves.

Harman claims, as we have seen, that Sturgeon's examples of moral explanations do not imply that moral properties themselves are causally effective. Earlier we saw that he can be interpreted to maintain that we never refer only to moral properties in explaining moral observations. It is

---

<sup>24</sup> Cf. Majors (2003), p. 134.

worth noticing that in making these claims, Harman might be misled by certain features of moral explanations. Above all, we provide such explanations on the assumption that the non-moral properties on which moral properties supervene have causal powers.<sup>25</sup> Moreover, it is reasonable to believe that non-moral properties are causally more fundamental than moral properties. To see this, consider the following example. Suppose someone explains the fact that a person performed certain actions by claiming that the person is good. However, imagine that it is objected that it is not quite clear how the goodness of the person explains these actions. The person who offered the explanation will then presumably try to answer this objection by providing an explanation in terms of the non-moral properties on which, to her mind, the goodness of the person supervenes. She might for example claim that the person in question performed the actions because she is friendly and generous. This indicates that the actions are causally explained by the non-moral properties on which the goodness of the person supervenes. It might also be taken to suggest that the causal powers ultimately are had by the non-moral properties, not the moral property. The reason is that the explanatory significance of the person's goodness is accounted for in terms of the causal powers of the non-moral properties. That is, what she says suggests that the person's goodness explains her actions because the non-moral properties on which her goodness supervenes causally explain these actions. Similar considerations apply to other kinds of moral explanations and other moral properties.

It is however important to notice what these features of moral explanations do *not* imply. First, they do not imply that moral properties are not *also* causally effective. We may hesitate to claim that moral properties have causal powers that are independent of the causal powers of the non-moral properties on which they supervene. However, they may have their causal powers in virtue of the causal powers of the non-moral properties on which they supervene. Moreover, even if they have their causal powers in virtue of the causal powers of the non-moral properties, their causal powers may differ from the causal powers of the non-moral properties in a certain way. (I will return to this possibility in section 7.) Second, even if moral properties do not have any causal powers at all and the causal powers only are possessed by the non-moral properties, this does not mean that the

---

<sup>25</sup> Cf. Audi (1993), pp. 100–103.

assumption that there are moral properties is not part of the best explanations. As we have seen, this assumption might be part of such explanations because moral properties strongly supervene on non-moral properties that have causal powers. Thus, Harman cannot appeal to the mentioned features of moral explanations to support his view.

Let us now turn to the second of Harman's arguments. This argument too rests on a conflation between the causal criterion and the best explanation criterion. As I read Harman, he claims that moral properties as characterised by moral epiphenomenalism are completely irrelevant to explanations even if they do not qualify as such according to Sturgeon's test because they are not causally effective themselves. In claiming this, he cannot have the best explanation criterion in mind, since entities on this criterion may be causally ineffectual and yet not completely irrelevant to explanations. Rather, he seems to have the causal criterion in mind. However, as he has abandoned this criterion, he cannot appeal to it.

There is a further reason why Harman's appeal to moral epiphenomenalism does not support his view. What Harman says about moral epiphenomenalism suggests that he believes that this view characterises moral properties in such a way that they are completely irrelevant to explanations. In Harman's vocabulary, this means that the assumption that there are such moral properties is not part of the best explanations, from which it, according to the best explanation criterion, follows that we are not justified in accepting their existence. But, as already has been suggested, it does not seem plausible to claim that moral properties do not exist on moral epiphenomenalism. Recall again that moral properties strongly supervene on non-moral properties. On moral epiphenomenalism this means that it is necessary that if something has a moral property, it has a set of non-moral properties—which is causally effective—such that the following holds: necessarily, if something has that set of non-moral properties, it has the moral property. Since the non-moral properties in question are causally effective and thus quite evidently exist, it seems hard to deny that the moral property that is necessitated by them also exists. There is a parallel here in mental epiphenomenalism, the view on which moral epiphenomenalism is modelled.<sup>26</sup> According to this view, mental properties

---

<sup>26</sup> However, it might be the case that, according to mental epiphenomenalism, mental properties do not *strongly* supervene on physical properties.



supervene on causally effective physical properties, but are not causally effective themselves. However, as far as I understand, this view is not generally taken to imply that mental properties do not exist. In any case, to deny the existence of epiphenomenal properties seems quite controversial. Robert Audi makes a similar point in relation to moral properties: ‘To be sure, if there are epiphenomena—items that have no casual power but appear only as effects—then being real is consistent with being inconsequential.’<sup>27</sup>

### 5. *A Dilemma for Harman’s Argument*

The analogy between moral properties and mental properties suggested by Harman’s appeal to moral epiphenomenalism raises an important issue. Harman believes, as we have seen, that moral properties do not fulfil the best explanation criterion unless they are reducible. Now, there are a number of other properties that are considered to be irreducible, among others mental properties and colours. It might then be asked which implications Harman’s argument has for these properties on the assumption that they are irreducible. As far as I see, Harman’s argument faces a dilemma.<sup>28</sup> The claim that irreducible moral properties do not fulfil the best explanation criterion seems to commit him to the same view as regards other irreducible properties. But it seems implausible to claim that we are not justified in believing in the existence of for instance mental properties and colours thus understood. In any event, such a view is quite controversial. This indicates that the criterion is too strong. Harman may then respond by maintaining that these properties do fulfil the criterion even if they are irreducible. But then it is difficult to see that he can maintain that irreducible moral properties do not fulfil the criterion as well. Harman’s argument seems in other worlds to face the following dilemma: either other irreducible properties, like irreducible moral properties, do not fulfil the best explanation criterion, in which case the criterion is too strong, or irreducible moral properties, like other irreducible properties, do fulfil the criterion. On both alternatives, Harman has failed to demonstrate that we are not justified in believing in the existence of irreducible moral properties. In order to avoid this dilemma, Harman has to make plausible that there is a

---

<sup>27</sup> Audi (1993), p. 98.

<sup>28</sup> For a related line of reasoning, see Sayre-McCord (1988), pp. 274–275. See also Pressler (1988), pp. 359–364, and Quinn (1986), pp. 537–539.

difference between irreducible moral properties and other irreducible properties which explains why we are justified in believing in the existence of the latter but not in the former. In the next section, I will consider Harman's proposal as to what such a difference might consist in.

In this context, it is suitable to respond to an objection that might be directed against my reasoning. I have argued that Harman is not successful in his attempt to show that irreducible moral properties do not fulfil the best explanation criterion. It might be argued that this view presumes a too weak conception of what provides the best explanations and that if we invoke a more demanding one, we will reach the result that irreducible moral properties do not fulfil this criterion. Brian Leiter suggests that whether an explanation is better than another is due to two factors: its consilience and its simplicity. An explanation is more consilient than another 'if it explains more classes of facts than the other does'.<sup>29</sup> Leiter argues that there are 'naturalistic' explanations in terms of e.g. evolutionary theory or psychoanalysis that are simpler and more consilient than moral explanations. He concludes that moral properties do not fulfil the best explanation criterion and that we consequently are not justified in believing in their existence.

In reply to this objection, it may first be pointed out that what criteria determine whether an explanation is better than another, and how these criteria should be interpreted, are controversial issues.<sup>30</sup> Any argument that rests on a particular understanding of these things is therefore bound to be contentious. In the absence of agreement on these matters, it may therefore be preferable to adhere to a conception of explanation whose implications are in line with common sense, even if it is not clearly specified. I think it is such a conception we have in mind when we take Sturgeon's examples to be evidence that moral properties explain various facts.

Moreover, it might be argued that Leiter does not succeed in showing that moral properties do not fulfil the best explanation criterion even if the standards are understood in the way he proposes.

Consider first Leiter's view that there are naturalistic explanations that are more consilient than moral explanations. It is important to be clear over how consilience should be understood to be relevant to Leiter's argument.

---

<sup>29</sup> Leiter (2001), pp. 80–81.

<sup>30</sup> Cf. Griffin (1996), p. 62.

As far as I understand, for an explanation E to be more consilient than an explanation E' in a way that provides support to Leiter's contention, it is not sufficient that E explains more classes of facts than E' merely in the sense that E applies to a larger number of facts than E'. It has also to be the case that E explains the facts that E' explains. The reason is that if the latter condition is not fulfilled, E' might have an advantage over E in terms of explanatory value by being able to explain facts that E cannot explain. In that case, it seems mistaken to claim that E is more consilient than E' in a way that justifies the claim that E is a better explanation than E'. Suppose that there are naturalistic explanations, e.g. psychoanalytic explanations, which explain a larger number of facts than moral explanations. However, it might be suspected that they are not able to explain the facts that the moral explanations explain. If that suspicion is correct, they do not seem to be more consilient than moral explanations in the relevant sense. Consequently, it does not follow that moral properties do not fulfil the best explanation criterion.

Moreover, we might question Leiter's reason for believing that moral explanations are less consilient than naturalistic explanations. The reason he gives for this claim is that naturalistic explanations explain both non-moral and moral facts whereas moral explanations merely explain moral facts.<sup>31</sup> This claim seems unfounded. As can be seen from the examples of moral explanations Sturgeon mentions, we refer to moral properties to explain various non-moral facts.<sup>32</sup> For example, a person's goodness can explain why she helps people in need, and the injustice of a society can explain why a revolt among its citizens broke out.

Consider next Leiter's view that there are naturalistic explanations that are simpler than moral explanations. Leiter's reason for this claim is that moral explanations involve the assumption that there are moral properties whereas naturalistic explanations do not. However, simplicity is a two-edged sword. An explanation that does not refer to moral properties is surely simpler in the mentioned respect. But in another respect it is

---

<sup>31</sup> Leiter (2001), p. 88. Crispin Wright argues in a similar way with reference to what he calls the 'wide cosmological role' of explanations: Wright (1992), pp. 191–199, and Wright (1996), pp. 14–15. For a criticism of naturalistic explanations in a moral context, see Sturgeon (1992), pp. 101–110.

<sup>32</sup> Cf. Shafer-Landau (2003), p. 103.

presumably less simple, since it has us to rewrite all our moral explanations in non-moral terms. Such explanations might be extremely complicated.

Finally, Leiter's line of reasoning faces the dilemma presented above. Suppose it is argued that irreducible moral properties do not fulfil the best explanation criterion because explanations that refer to such properties are less consistent and simple than explanations that do not. On this view, there is reason to believe that other properties that are thought to be irreducible, such as mental properties and colours, fail to fulfil the best explanation criterion for the same reason. However, it seems untenable to claim that these properties do not exist. In any case, such a conclusion is quite controversial. Suppose it is replied that such properties do fulfil the best explanation criterion because explanations which refer to them are more consistent and simpler than alternative explanations. However, in that case it seems difficult to avoid the corresponding conclusion concerning irreducible moral properties.

We may now summarise this section and the previous one. Harman argues that moral properties fulfil the best explanation criterion only on condition that they are reducible. However, we have found reasons to maintain that moral properties fulfil the criterion even if they are irreducible. Moreover, the dilemma above suggests that even if irreducible moral properties do not fulfil the criterion, this should not make us claim that we are not justified in believing in their existence; rather, it indicates that the criterion is too strong.

## 6. *The Pragmatic Criterion*

In the last section, I argued that Harman's argument faces a dilemma regarding the best explanation criterion unless he can show that there is a relevant difference between irreducible moral properties and other irreducible properties. Harman can be understood to address this worry with regard to colours. He seems willing to admit that colours might be irreducible.<sup>33</sup> However, he still thinks there would be an important difference between colours thus understood and irreducible moral properties:

[W]e will still *sometimes* refer to the actual colors of objects in explaining color perception, if only for the sake of simplicity. [- - -] We will continue to believe that

---

<sup>33</sup> Harman (1977), p. 22.

objects have colors because we will continue to refer to the actual colors of objects in the explanations that we will in practice give. A similar point does not seem to hold for moral facts. There does not ever seem to be, even in practice, any point to explaining someone's moral observations by appeal to what is actually right or wrong, just or unjust, good or bad. [- - -] So, the reasons we have for supposing that there are facts about colors do not correspond to reasons for thinking that there are moral facts.<sup>34</sup>

It seems reasonable to assume that Harman appeals to the following criterion:

*The pragmatic criterion:* We are justified in believing that a certain entity exists only if it is practically useful to refer to it in explanations of our observations with regard to it.<sup>35</sup>

For the criterion to be relevant for the discussion in the last section, we might formulate the following argument on Harman's behalf. Even if colours are irreducible, it is practically useful to refer to them in explanations of colour observations. However, it is not practically useful to refer to irreducible moral properties in explanations of moral observations. Therefore, we are not justified in believing in the existence of such properties according to the pragmatic criterion.

In response to this argument, it should be pointed out that in the same way as it may be practically useful to refer to colours, it may be practically useful to refer to irreducible moral properties.<sup>36</sup> There might be a number of reasons why this is the case. Harman proposes that explanations in terms of colours are simpler, and therefore more practical, than explanations in terms of physical properties. Correspondingly, explanations in terms of moral properties might be simpler than explanations in terms of the underlying non-moral properties. This would especially be the case if moral properties supervene on a number of sets of non-moral properties or if such a set of non-moral properties consists of long and complicated conjunctions of non-moral properties. It is therefore difficult to see that Harman succeeds in showing that there is a difference between irreducible

---

<sup>34</sup> Harman (1977), p. 22. Harman says that reductions of colours might be 'complex, vague, and difficult (probably impossible) to specify' (Harman (1977), p. 22). As I interpret Harman, he leaves room for the possibility that colours are irreducible in the sense of 'reduction' I use here.

<sup>35</sup> It might be objected that Harman does not intend to put forward a necessary condition. But it is doubtful whether it would strengthen Harman's argument if he instead is understood to propose a sufficient condition. For one thing, it does not seem plausible to assume that the fact that it is practically useful to refer to an entity in explanations is enough to establish that we are justified to believe in its existence.

<sup>36</sup> Cf. Brink (1989), p. 192; Sayre-McCord (1988), p. 275, and Sturgeon (1984), p. 58.

moral properties and colours which justifies the view that the former do not fulfil the pragmatic criterion whereas the latter do.

In this context, it is appropriate to draw attention to some ways in which irreducible moral properties might play important roles in explanations of various facts.<sup>37</sup>

For instance, it seems reasonable to argue that explanations in terms of irreducible moral properties might provide information on a more suitable level of generality than explanations merely in non-moral terms. Consider the following example of how this might be the case. Suppose the property of being good supervenes on a number of sets of non-moral properties. Suppose further that persons who are good regularly are causally connected to certain facts. For instance, it might be the case that they typically perform certain actions and evoke certain reactions. Assume now that it is asked about a certain person why she performed a certain action or why people responded to her action in a certain way. In some situations the answer 'She's a good person' might provide a better explanation than an answer in terms of the set of non-moral properties which the goodness of the person supervenes on in the particular case at issue. The reason is that the person would have acted roughly in the same way and would have evoked roughly the same reactions if she instead had instantiated some of the other sets of non-moral properties that goodness supervenes on. It might then be fairly unimportant which particular set of non-moral properties the good person instantiates in this particular case; such information is simply superfluous. Thus, explanations in terms of an irreducible moral property might provide information on a more suitable level of generality than explanations in terms of the various sets of non-moral properties which the moral property supervenes on.

This reasoning suggests that explanations in terms of irreducible moral properties may have a unique explanatory function. As the example just mentioned indicates, by referring to a moral property it is possible to provide one single explanation of a whole collection of facts. No single explanation only in terms of the sets of non-moral properties that the moral property supervenes on would be able to do this, since each such explanation would refer to different sets of non-moral properties. Thus, to have one

---

<sup>37</sup> For related arguments, see Brink (1989), pp. 194–195; Majors (2003), pp. 139–140, Railton (1998), pp. 178–179, and Sayre-McCord (1988), pp. 275–276. For criticism of these arguments, see Leiter (2001), pp. 93–101, and Zimmerman (1984), pp. 85–87.

single explanation of these facts, we have to adhere to an explanation stated in terms of an irreducible moral property; explanations in terms of non-moral properties cannot have this function.

Moreover, moral explanations may allow us to identify regularities that are difficult or impossible to identify merely in non-moral terms. Suppose again that the property of being good supervenes on a number of sets of non-moral properties. Suppose further that persons who are good generally are causally connected to certain facts; they typically perform certain actions and evoke certain reactions. It may be the case that the regularities that hold between people who are good and the mentioned facts are difficult or impossible to identify in terms of the non-moral properties which goodness supervenes on. One reason might be that these relations are immensely complicated because goodness supervenes on a great number of sets of non-moral properties. However, these regularities might be identifiable in terms of goodness. Like the previous observation, this observation indicates that explanations in terms of an irreducible moral property may have a distinct explanatory function that explanations in terms of non-moral properties cannot fulfil.

There are consequently reasons to believe that irreducible moral properties may have important roles to play in explanations of various facts. It therefore seems plausible to think that if irreducible moral properties figure in these explanations, the assumption that there are such properties is part of the best explanations. It is important to observe that the mentioned explanations do not presuppose that moral properties have causal powers themselves. It might be the case that the moral properties appealed to in these explanations do not have any causal powers, but that the causal powers merely accrue to the non-moral properties that the moral properties supervene on. This means that the assumption that there are irreducible moral properties might be part of the best explanations even if such properties do not have any causal powers. Hence, irreducible moral properties might pass the best explanation criterion without being causally efficacious themselves. Although I used goodness in these examples, I think my claims are valid for other moral properties as well.

### 7. Three Conceptions of the Causal Powers of Moral Properties

In the previous sections, I argued that there are reasons to believe that irreducible moral properties fulfil the three explanatory criteria Harman appeals to. The first two of these criteria concern, broadly understood, the causal explanatory role of moral properties. In this section, I will inquire how realists may conceive of the causal powers of moral properties.

The discussion in section 4 suggests that there are three alternative conceptions of the causal powers of moral properties:<sup>38</sup> **(i)** A moral property M has causal powers that differ from the causal powers of the non-moral properties it supervenes on.<sup>39</sup> **(ii)** A moral property M has the same causal powers as the non-moral properties it supervenes on. **(iii)** A moral property M does not have any causal powers; the pertinent causal powers are had by the non-moral properties it supervenes on.<sup>40</sup> In the preceding sections, we saw that there are reasons to maintain that, on any of these alternatives, irreducible moral properties fulfil the best explanation criterion. On the two first alternatives, but not the third, such moral properties also fulfil the causal criterion.

In what follows, I will consider the three alternatives in turn. In doing so, I will consider which assumptions would lead advocates of realism to adopt these alternatives and whether these alternative pose any difficulties for this view. But first an important caveat is in place. The three alternatives raise fundamental metaphysical issues that I am not able to deal with here. This means that my discussion of moral causation is incomplete; especially, there might be issues concerning the nature of causation and the nature of properties that are relevant to the plausibility of these alternatives but which I fail to take into consideration.

*Alternative (i).* According to this alternative, the causal powers of a moral property M differ from the causal powers of the non-moral properties

---

<sup>38</sup> Alternatives **(i)** – **(iii)** have counterparts in various views about the causal powers of mental properties. Quite generally, much of the discussion on mental causation is relevant for the question whether moral properties have causal powers and, if they do, how this should be understood. Unfortunately, I am not able to discuss the relation between moral and mental causation here. For relevant discussions, see e.g. Brink (1989), pp. 195–196; Majors (2003), pp. 135–138, and Shafer-Landau (2003), pp. 108–110.

<sup>39</sup> Different alternatives may hold for different moral properties. For the sake of simplicity, I assume here that the same alternative applies to all moral properties. Moreover, I assume that a moral property M supervenes on more than one set of non-moral properties.

<sup>40</sup> This view corresponds to what Harman calls ‘moral epiphenomenalism’.



it supervenes on. As far as I see, realists can defend this alternative only in a quite weak version. There are basically three assumptions realists can make that would generate this view.

The first assumption is that moral properties have causal powers. As we saw in section 3, this view is supported by our conception of moral properties as it manifests itself in the moral explanations we provide of various facts.

The second assumption is that in so far as moral properties have any causal powers, they have such powers in virtue of the causal powers of the non-moral properties on which they supervene. As was argued in section 4, this view finds support in the way in which we justify claims to the effect that moral properties causally explain facts; for instance, we justify the claim that a person's goodness causes her to perform certain actions by citing the non-moral properties on which we take her goodness to supervene. More generally, this view finds support in the layered conception of reality, according to which the causal powers of one level of reality are had in virtue of the causal powers of the level next 'below'.<sup>41</sup> Another way of formulating this conception is that moral properties inherit their causal powers from the non-moral properties on which they supervene.

In the previous chapters, I argued that realists should understand the dependence of moral properties on non-moral ones in terms of the realist formula. Since I in this section will discuss how realists should conceive of the causal powers of moral properties, I will make use of this notion. In accordance with the second assumption, realists may thus suggest that if a moral property M has any causal powers, it has the causal powers of the sets of non-moral properties that make objects have M, where 'make' is understood in line with the realist formula.<sup>42</sup> One explanation of this relation between causal powers is that token identity holds between instances of M and instances of such sets, i.e. that each instance of M is identical with a certain instance of a set of non-moral properties that makes objects have M.<sup>43</sup> To apply a point often made in the philosophy of mind, this is compatible with M not being identical with a certain set of non-

---

<sup>41</sup> See e.g. Kim (1993 (1993)), pp. 337–339.

<sup>42</sup> These sets of non-moral properties might in turn have *their* causal powers in virtue of the causal powers of the properties they supervene on. This is an aspect of the layered conception of reality.

<sup>43</sup> Cf. Kim (1993 (1993)), p. 355.

moral properties.<sup>44</sup> (It might seem strange to say that a set of properties has causal powers; what have causal powers are presumably properties or their instances. However, for the sake of simplicity, I will stick to the terminology adopted in previous chapters.)

The third assumption is that the sets of non-moral properties which make objects have M have different causal powers. That is, at least two of the sets which make objects have M do not have the same causal powers.

We can now see how a certain version of alternative (i) is obtained from these assumptions. According to the two first assumptions, a moral property M has causal powers and it has the causal powers of the sets of non-moral properties which make objects have this moral property. According to the third assumption, these sets have different causal powers. These assumptions imply that M does not have the same causal powers as any particular one of these sets. Notice next that these assumptions imply that M has causal powers that outreach the causal powers of each individual set of non-moral properties which makes objects have this moral property. The moral property M has the causal powers of the various sets of this kind. Since these sets have different causal powers, M has causal powers that exceed the causal powers of each such set. It might then be argued that there is a certain sense in which the causal powers of M differ from the causal powers of the non-moral properties it supervenes on. Admittedly, this version of alternative (i) is quite weak, since it does not mean that M has any causal powers that are not had by any of these sets of non-moral properties. Rather, it has the causal powers of all these sets taken together. In other words, M does not have any novel causal powers as compared with the causal powers of these sets of non-moral properties. As far as I see, it is difficult for realists to maintain a stronger version of this alternative, granted that they accept the second assumption. However, I want to leave it open whether a stronger version of (i) is feasible.<sup>45</sup>

We might now ask if adopting alternative (i) would pose any difficulties for realism. It might be argued that according to a certain notion of properties, the present version of this alternative is too weak for realism.

---

<sup>44</sup> In case the set of non-moral properties in question consists of a conjunction of properties, we might assume that properties are closed under conjunction and hence that such a set comprises one property.

<sup>45</sup> Cf. Pereboom and Kornblith (1990), pp. 128–132, 142–144.

On this notion, proposed by Jaegwon Kim, a property is 'real' only if it has causal powers that are not had by any of the properties it supervenes on.<sup>46</sup>

The notion of a 'real' property appealed to in the argument cannot be dealt with in the present context, but it is reasonable to assume that the sense in which Kim uses 'real' is not relevant for whether a certain property exists. There are various views to the effect that properties of a certain kind have their causal powers inherited from the properties they supervene on and hence do not have any novel causal powers as compared with these properties. However, our responses to these views do not seem to be that they imply that these properties do not exist and are not 'real' in that sense. On the notion that a property is not 'real' unless it has novel causal powers, there is presumably a number of properties that would not qualify as 'real'; moral properties would consequently be in good company.<sup>47</sup> Rather than concerning the existence of properties, the sense of 'real' Kim appeals to seems to concern whether the properties in question belong to reality's fundamental causal constituents. However, that moral properties are not 'real' in that sense should readily be accepted by realists.

*Alternative (ii).* According to this alternative, a moral property M has the same causal powers as the non-moral properties it supervenes on.

Notice first that it is hard to deny the first and second assumption mentioned above. As we have seen, the first assumption finds support in our notion of moral properties as it reveals itself in our moral explanations. The second assumption is difficult to deny, for doing so would be to deny the way in which we justify moral explanations. Moreover, it would be to deny the layered conception of the world, a picture with considerable intuitive appeal. However, suppose realists deny the third assumption. That is, suppose it is thought that the sets of non-moral properties which make objects have a moral property M have the same causal powers. In that case, M has the same causal powers as each such set. Then realists should adopt alternative (ii).

Would this alternative pose a problem for realism? According to one conception of properties, sameness of causal powers is constitutive of property identity.<sup>48</sup> On this notion, it might be thought that if a property A

---

<sup>46</sup> See e.g. Kim (1993 (1993)), p. 350. Kim refers to this principle as 'Alexander's dictum'.

<sup>47</sup> Cf. Shafer-Landau (2003), p. 110.

<sup>48</sup> See e.g. Mumford (1998), pp. 121–125. Cf. Armstrong (1978), pp. 43–47.

and a property B have the same causal powers, they are identical. If this is correct, it might be suspected that alternative (ii) is incompatible with realism.<sup>49</sup>

In response to this argument, it should first be pointed out that it is controversial whether sameness of causal powers is constitutive of property identity. Second, even if sameness of causal powers is thus constitutive, it does not comprise a sufficient condition for property identity.<sup>50</sup> At least one more condition has to be fulfilled for a property A to be identical with property B: A has to be necessary co-extensive with B, so that a necessary biconditional involving A and B holds.<sup>51</sup> Something that might be problematic for realism is if a moral property M and the sets of non-moral properties that make objects have this moral property have the same causal powers *and* there is a necessary biconditional involving M and a disjunctive property  $D_M$  consisting of these sets of non-moral properties. The latter would be the case only on condition that disjunction is a proper way of forming properties. However, in the preceding chapter we saw that there are reasons to deny this view. Moreover, it is controversial whether necessary co-extensiveness and sameness of causal powers constitute necessary and jointly sufficient conditions for property identity.<sup>52</sup> I think we have to accept that, once again, we have reached a point where the discussion of realism raises fundamental metaphysical questions that cannot be dealt with here. Consequently, whether considerations in relation to the causal powers of moral properties provide reasons to believe that moral properties are identical with non-moral properties and that realism therefore is mistaken cannot be settled in the present thesis.

*Alternative (iii).* According to this alternative, a moral property M does not have any causal powers; the causal powers associated with M are merely had by the non-moral properties it supervenes on. As this alternative

---

<sup>49</sup> A similar argument might be directed against realism if it is combined with the weak version of alternative (i) described above, since a moral property M according to this version has the same causal powers as the sets of non-moral properties which make objects have this moral property, taken together.

<sup>50</sup> See e.g. Achinstein (1974), pp. 273–274; Clarke (1999), p. 300–301; Heil (1999), pp. 192–193, and Macdonald (1989), pp. 223–224.

<sup>51</sup> For a proof to the effect that sameness in causal powers does not imply necessary co-extensiveness, see Achinstein (1974), pp. 273–274.

<sup>52</sup> According to one conception of properties, they have a qualitative aspect as well as a dispositional or causal aspect; see e.g. Martin (1997), pp. 213–217.

denies the first assumption mentioned above, there is reason for realists to oppose it.

However, there is a certain argument that would provide realists with a reason to adopt alternative (iii). The argument I have in mind is a version of an argument originally formulated by Jaegwon Kim in relation to mental causation.<sup>53</sup> However, it might with some revisions be mapped over to the moral area. The question posed in the original argument is how a mental property and the physical property it supervenes on in a particular instance of it are causally related to a certain result. In the moral case, and formulated in terms of the realist formula, the question is how a moral property M and the set of non-moral properties G which makes an object have this moral property in a particular instance of it are causally related to a certain result. In the argument thus understood, two options are considered. The *first option* is that M and G have causal powers such that M and G are *jointly, but not separately*, sufficient for the result. The *second option* is that M and G have causal powers such that M and G *each, separately*, are sufficient for the result. Now, both these options give rise to difficulties according to the argument. The first option seems to violate the second assumption above. Since both M and G are needed to bring about the result on this option, M seems to have causal powers that are not derived from the sets of non-moral properties which make objects have this moral property. The second option is taken to imply that the result in question is causally overdetermined, a consequence Kim thinks should be avoided, at least when it comes to mental causation. The two options rest on the assumption that both a moral property M and the sets of non-moral properties that make objects have this moral property have causal powers. Since both of these options are considered to be untenable, the argument may be taken to suggest that M does not have any causal powers, but that the causal work is done solely by the non-moral properties.

As mentioned above, there is reason for realists to oppose alternative (iii). We should therefore inquire whether alternatives (i) and (ii) can evade

---

<sup>53</sup> See e.g. Kim (1993 (1993)), pp. 350–353. Kim does not consider this argument as decisive; he regards it rather as an argument for mental epiphenomenalism that has to be responded to by those who hold that mental properties have causal powers. Moral analogues to this argument are considered by Majors (2003), pp. 135–138, and Shafer-Landau (2003), pp. 106–110.

the argument. It might be argued that they can, and the reason is that both recognise the second assumption mentioned above.

Let us consider whether alternative (i) can avoid the argument. Notice first that the version of this alternative considered above does not have the difficulty with the first option. Since it recognises the second assumption, it says that a moral property M does not have any causal powers that are not had by any of the sets of non-moral properties which make objects have this moral property. Thus, M and a set of non-moral properties G do not have different causal powers such that M and G are jointly, but not separately, sufficient for a particular result. It might further be argued that alternative (i) does not have the problem with the second option. Again, according to the version of alternative (i) under consideration, M does not have any causal powers that are not had by any of the sets of non-moral properties which make objects have this moral property. This suggests that M and G do not have causal powers such that G with *its* causal powers is sufficient for a particular result and that M with *its* additional causal powers is sufficient for the same result. There are thus not, strictly speaking, two separate causal explanations of why a particular result occurs; rather, there is only one causal explanation that can be given either in terms of M or in terms of G.<sup>54</sup> If this is correct, it means that alternative (i) does not lead to overdetermination, at least not if this alternative is understood in the way described earlier.<sup>55</sup> A similar line of reasoning holds for alternative (ii).

It can further be argued that even if alternative (i) or (ii) would result in overdetermination, this would not be of a kind that is especially

---

<sup>54</sup> If I understand Kim correctly, he provides a similar account of mental causation in reply to the argument in relation to mental properties; see e.g. Kim (1993 (1993)), pp. 352–353. However, Kim seems to think that mental properties on such a view would not be ‘real’ properties, i.e. properties with novel causal powers. He also suggests that an analogous account holds for the causal powers of moral properties: Kim (1983), pp. 53–54. As regards mental causation, a similar view is proposed by e.g. Clarke (1999), pp. 304–313.

<sup>55</sup> It might be objected that when M is instantiated, it has also other causal powers than those had by the set of non-moral properties which makes an object have this moral property in that particular instance of it. The reason would be that M also has the causal powers of the other sets of non-moral properties which make objects have this moral property. It might then be suggested that these causal powers, together with the causal powers of the set which makes an object have M in the particular instance at issue, results in overdetermination. It is difficult to answer this objection in the present context, since it raises questions that I am not able to deal with here. Especially, it brings up the question whether M, in each of its instances, carries with it all its causal powers, i.e. all the causal powers it has in virtue of the various sets of non-moral properties that make objects have M. Some authors seem to take this assumption for granted; see e.g. Clarke (1999), p. 310.

problematic. If I understand Kim correctly, in the cases he considers overdetermination arises because a mental property is presumed to have causal powers that are not had by any of the physical properties it supervenes on. That is, in the cases he considers overdetermination arises because it is presumed that a mental property has novel causal powers as compared with the causal powers of the physical properties it supervenes on.<sup>56</sup> Kim believes that the view that mental properties have novel causal powers is associated with serious difficulties because it threatens the ‘causal closure of the physical domain’.<sup>57</sup> The corresponding view in relation to moral properties—that a moral property has novel causal powers as compared with the non-moral properties it supervenes on—would be problematic because it seems difficult to reconcile with the second assumption mentioned above. However, if overdetermination arises on alternative (i) or (ii), it would presumably not be because a moral property has novel causal powers. The reason is that both these alternatives acknowledge the second assumption—at least the versions of these alternatives I have considered here. This may be taken to suggest that if overdetermination arises on either of these alternatives, it would be of a kind that is fairly innocent.

There is a further reason to believe that it would not be particularly problematic for alternatives (i) and (ii) if they cannot avoid overdetermination. Moral properties are only one kind of properties among many that are thought to be supervenient on some underlying properties. Other examples are mental properties, colours and properties belonging to various special sciences, such as biology and sociology. Moreover, there are reasons to believe that these properties cannot be reduced to the underlying properties. Now, analogous arguments to the one above could be proposed for each such area; indeed, as mentioned above, the argument was originally presented in relation to mental causation. This suggests that, as regards all

---

<sup>56</sup> In such cases, a certain caused result has a sufficient causal explanation both in terms of the novel causal powers of a mental property and in terms of the causal powers of the physical properties which the mental property supervenes on.

<sup>57</sup> This is, in brief, the view that each physical fact has a sufficient causal explanation merely in terms of the causal powers of physical properties. If mental properties have novel causal powers as compared with the causal powers of the physical properties they supervene on, this view is threatened. The reason is that it would mean that a physical fact might have a sufficient causal explanation in terms of the causal powers of a mental property where these causal powers are not the same as the causal powers of the physical properties the mental property supervenes on.

these areas, we would have to choose between accepting overdetermination and accepting that the supervenient properties lack causal powers. Whichever of these alternatives holds for moral properties, it will consequently hold for properties in other areas as well.<sup>58</sup> Moreover, of the two alternatives, overdetermination seems preferable. For example, it seems more reasonable to maintain that both biological properties and the physical properties they supervene on have causal powers than to hold that biological properties do not have any causal powers at all, even if the first alternative would imply overdetermination.

Thus, there are reasons to believe that realists are not forced to adopt alternative (iii). Suppose, however, that this alternative is correct. Would this make realism vulnerable to any objections, besides that it would not be able to ascribe causal powers to moral properties? It might be argued that on the view that moral properties do not have any causal powers, we are not justified in believing in their existence. However, as was observed above, this notion seems too strong. In previous sections, we have also seen that there are reasons to believe that moral properties may exist even if they do not have any causal powers themselves.

As we have seen, there is reason for realists to oppose alternative (iii), since it implies that moral properties do not have any causal powers. However, there is one important reason why it might be tempting for realists to support this alternative. According to this alternative, a moral property M and what makes objects have this moral property do not have the same causal powers. Now, if a property A and a property B have different causal powers, A and B are evidently not identical. Accordingly, if it would turn out to be difficult to combine alternative (i) or (ii) with the view that moral properties are not identical with non-moral properties, alternative (iii) might become attractive to realists.

## 8. *Concluding Remarks*

Harman thinks that for us to be justified in believing in the existence of a certain entity, it has to fulfil an explanatory criterion. He proposes two main criteria of this kind, the causal criterion and the best explanation criterion, and argues that moral properties do not fulfil either of them, at least not if

---

<sup>58</sup> For basically the same point, see e.g. Majors (2003), pp. 136–137; Pereboom and Kornblith (1990), pp. 142–143, and Shafer-Landau (2003), pp. 109–110. Majors and Shafer-Landau accept overdetermination in the moral case as well in others.



they are irreducible. He therefore concludes that we are not justified in believing in the existence of such properties. However, he also admits that the causal criterion seems to be too strong. In this chapter, I have argued with Sturgeon that, according to our notion of moral properties as it manifests itself in our moral explanations, moral properties causally explain various facts. There are thus *prima facie* reasons to believe that moral properties are causally explanatory, and this view should not be abandoned in the absence of strong arguments. However, I have maintained that Harman fails to supply such arguments. Thus, it can be maintained that moral properties fulfil the causal criterion. Moreover, I have argued that there are reasons to believe that moral properties fulfil the best explanation criterion. In particular, I have maintained that Harman is not successful in his attempt to show that moral properties do not fulfil this criterion unless they are reducible to non-moral properties. Realists can consequently uphold the view that moral properties, as they understand them, exist. We have also found that in case irreducible moral properties would not fulfil the best explanation criterion, this should not make us draw the conclusion that such properties do not exist; rather, it indicates that the criterion is too strong. Moreover, I have argued that realists may conceive of the causal powers of moral properties according to three alternatives. However, it should be stressed that the discussion of causal powers of moral properties is incomplete, since it raises fundamental metaphysical issues that I have not discussed. As we have seen, this limitation is important because these issues might be relevant to whether the causal role of moral properties is compatible with realism.

In chapter 1, I mentioned different senses in which properties may be natural properties. According to one of these, properties are natural in so far as they are part of nature by being involved in its causal relations. In this chapter, we have found that realists can maintain that moral properties are natural properties in this sense. Moreover, the discussion in this chapter suggests that the non-moral properties which make objects have moral properties are natural properties in this sense.



## References

- Achinstein, Peter (1974), 'The Identity of Properties', *American Philosophical Quarterly*, Vol. 11.
- Adams, Robert Merrihew (1981 (1973)), 'A Modified Divine Command Theory of Ethical Wrongness', *Divine Commands and Morality*, ed. by Paul Helm, Oxford: Oxford University Press.
- Adams, Robert Merrihew (1981 (1979)), 'Divine Command Metaethics as Necessary a Posteriori', *Divine Commands and Morality*, ed. by Paul Helm, Oxford: Oxford University Press.
- Alm, David (2004), 'Atomism about Value', *Australasian Journal of Philosophy*, Vol. 82.
- Alston, William P. (1996), *A Realist Conception of Truth*, Ithaca: Cornell University Press.
- Armstrong, D. M. (1978), *A Theory of Universals*, Vol. II, Cambridge: Cambridge University Press.
- Arrington, Robert L. (1989), *Rationalism, Realism, and Relativism*, Ithaca: Cornell University Press.
- Audi, Robert (1993), 'Ethical Naturalism and the Explanatory Power of Moral Concepts', *Naturalism. A Critical Appraisal*, ed. by Steven J Wagner and Richard Warner, Notre Dame: University of Notre Dame Press.
- Audi, Robert (1996), 'Intuitionism, Pluralism, and the Foundations of Ethics', *Moral Knowledge? New Readings in Moral Epistemology*, ed. by Walter Sinnott-Armstrong and Mark Timmons, Oxford: Oxford University Press.
- Audi, Robert (1997), 'Moral Judgement and Reasons for Action', *Ethics and Practical Reason*, ed. by Garrett Cullity and Berys Gaut, Oxford: Clarendon Press.
- Ayer, A. J. (1952), *Language, Truth and Logic*, New York: Dover Publications.
- Bailey, Andrew (1998/99), 'Supervenience and Physicalism', *Synthese*, Vol. 117.
- Baldwin, Thomas (1985), 'Ethical Non-naturalism', *Exercises in Analysis*, ed. by Ian Hacking, Cambridge: Cambridge University Press.
- Baldwin, Thomas (1990), *G. E. Moore*, London: Routledge.
- Baldwin, Thomas (2003), 'The Indefinability of Good', *The Journal of Value Inquiry*, Vol. 37.
- Ball, Stephen W. (1988), 'Reductionism in Ethics and Science: A Contemporary Look at G. E. Moore's Open Question Argument', *American Philosophical Quarterly*, Vol. 25.

- Ball, Stephen W. (1989), 'Facts, Values, and Normative Supervenience', *Philosophical Studies*, Vol. 55.
- Ball, Stephen W. (1991), 'Linguistic Intuitions and Varieties of Ethical Naturalism', *Philosophy and Phenomenological Research*, Vol. LI.
- Barker, Stephen J. (2000), 'Is Value Content a Component of Conventional Implicature?', *Analysis*, Vol. 60.
- Barnett, David (2002), 'Against *A Posteriori* Moral Naturalism', *Philosophical Studies*, Vol. 107.
- Bennigson, Thomas (1996), 'Irresolvable Disagreement and the Case against Moral Realism', *The Southern Journal of Philosophy*, Vol. XXXIV.
- Björnsson, Gunnar (1998), *Moral Internalism. An Essay in Moral Psychology*, Stockholm: Preprint from Stockholm University.
- Blackburn, Simon (1984), *Spreading the Word*, Oxford: Oxford University Press.
- Blackburn, Simon (1993), *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Blackburn, Simon (1993 (1971)), 'Moral Realism', *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Blackburn, Simon (1993 (1985)), 'Supervenience Revisited', *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Blackburn, Simon (1993 (1988)), 'How to be an Ethical Anti-Realist', *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Blackburn, Simon (1993 (1990)), 'Just Cases', *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Blackburn, Simon (1998), *Ruling Passions*, Oxford: Clarendon Press.
- Bovens, Luc and Dalia Drai (1999), 'Supervenience and Moral Realism', *Philosophia*, Vol. 27.
- Boyd, Richard N. (1984), 'The Current Status of Scientific Realism', *Scientific Realism*, ed. by J. Leplin, Berkeley: University of California Press.
- Boyd, Richard N. (1988) 'How to Be a Moral Realist', *Essays on Moral Realism*, ed. by Geoffrey Sayre-McCord, Ithaca: Cornell University Press.
- Brandt, Richard B. (1959), *Ethical Theory*, Englewood Cliffs: Prentice-Hall.
- Brandt, Richard B. (1979), *A Theory of the Good and the Right*, Oxford: Clarendon Press.
- Brandt, Richard B. (1996), *Facts, Values, and Morality*, Cambridge: Cambridge University Press.
- Brink, David O. (1989), *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.
- Brink, David O. (1992), 'A Puzzle About the Rational Authority of Morality', *Philosophical Perspective*, 6, Ethics.
- Brink, David O. (1997), 'Moral Motivation', *Ethics*, Vol. 108.

- Brink, David O. (2000), 'Objectivity and Dialectical Methods in Ethics', *Inquiry*, Vol. 42.
- Brink, David O. (2001), 'Realism, Naturalism and Moral Semantics', *Social Philosophy and Policy*, Vol. 18.
- Broad, C. D. (1971), *Broad's Critical Essays in Moral Philosophy*, ed. by David R. Cheney, London: George Allen & Unwin.
- Broad, C. D. (1971 (1934)), 'Is "Goodness" a Name of a Simple Non-natural Quality?', *Broad's Critical Essays in Moral Philosophy*, ed. by David R. Cheney, London: George Allen & Unwin.
- Brueckner, Anthony (2002), 'Blackburn's Modal Argument Against Moral Realism', *Theoria*, Vol. LXVIII.
- Butchvarov, Panayot (1988), 'Realism in Ethics', *Midwest Studies in Philosophy*, Vol. XII, Realism and Antirealism.
- Cargile, James (1989), 'What is a Natural Property?', *Philosophy*, Vol. 64.
- Clapp, Lenny (2001), 'Disjunctive Properties: Multiple Realizations', *The Journal of Philosophy*, Vol. XCVIII.
- Clarke, R. (1999), 'Nonreductive Physicalism and the Causal Powers of the Mental', *Erkenntnis*, Vol. 51.
- Copp, David (1995), 'Moral Obligation and Moral Motivation', *On the Relevance of Metaethics. New Essays On Metaethics*, ed. by Jocelyne Couture and Kai Nielsen, *Canadian Journal of Philosophy*, Supplementary Volume 21.
- Copp, David (1997), 'Belief, Reason, and Motivation: Michael Smith's The Moral Problem', *Ethics*, Vol. 108.
- Copp, David (2000), 'Milk, Honey, and the Good Life on Moral Twin Earth', *Synthese*, Vol. 124.
- Copp, David (2001), 'Realist-Expressivism: A Neglected Option for Moral Realism', *Social Philosophy and Policy*, Vol. 18.
- Copp, David (2003), 'Why Naturalism?', *Ethical Theory and Moral Practice*, Vol. 6.
- Cox, H. H. (1970), 'Warnock on Moore', *Mind*, Vol. LXXIX.
- Craig, Edward (1998), 'Realism and Antirealism', *Routledge Encyclopedia of Philosophy*, ed. by Edward Craig, London: Routledge.
- Crisp, Roger (2000), 'Particularizing Particularism', *Moral Particularism*, ed. by Brad Hooker and Margaret Olivia Little, Oxford: Oxford University Press.
- Cullity, Garrett and Berys Gaut (1997), 'Introduction', *Ethics and Practical Reason*, ed. by Garrett Cullity and Berys Gaut, Oxford: Clarendon Press.
- Cuneo, Terence D. (1999), 'An Externalist Solution to the "Moral Problem"', *Philosophy and Phenomenological Research*, Vol. LIX.
- Cuneo, Terence D. (2002), 'Reconciling Realism with Humeanism', *The Australasian Journal of Philosophy*, Vol. 80.

- Dancy, Jonathan (1981), 'On Moral Properties', *Mind*, Vol. XC.
- Dancy, Jonathan (1986), 'Two Conceptions of Moral Realism', *Proceedings of the Aristotelian Society*, Supplementary Volume LX.
- Dancy, Jonathan (1993), *Moral Reasons*, Oxford: Blackwell.
- Dancy, Jonathan (1998), 'Moral Realism', *Routledge Encyclopedia of Philosophy*, ed. by Edward Craig, London: Routledge.
- Dancy, Jonathan (1999a), 'On the Logical and Moral Adequacy of Particularism', *Theoria*, Vol. LXV.
- Dancy, Jonathan (1999b), 'Defending Particularism', *Metaphilosophy*, Vol. LLC.
- Dancy, Jonathan (2000), 'Should We Pass the Buck?', *Philosophy, the Good, the True and the Beautiful*, ed. by Anthony O'Hear, Cambridge: Cambridge University Press.
- Danielsson, Sven (2001), 'The Supervenience of Intrinsic Value', *Omnium-gatherum. Philosophical essays dedicated to Jan Österberg on the occasion of his sixtieth birthday*, ed. by Erik Carlson and Rysiek Sliwinski, Uppsala: Uppsala Philosophical Studies.
- Darwall, Stephen (1990), 'Autonomist Internalism and the Justification of Morals', *Noûs*, Vol. 24.
- Darwall, Stephen (2003), *Welfare and Rational Care*, Princeton: Princeton University Press.
- Darwall, Stephen, Allan Gibbard and Peter Railton (1992), 'Toward Fin de siècle Ethics: Some Trends', *The Philosophical Review*, Vol. 101.
- Depaul, Michael R. (1987), 'Supervenience and Moral Dependence', *Philosophical Studies*, Vol. 51.
- Devitt, Michael (1991), *Realism and Truth*, Oxford: Blackwell.
- Dreier, James (1992), 'The Supervenience Argument against Moral Realism', *The Southern Journal of Philosophy*, Vol. XXX.
- Dreier, James (2000), 'Dispositions and Fetishes: Externalist Models of Moral Motivation', *Philosophy and Phenomenological Research*, Vol. LXI.
- Dummett, Michael (1982), 'Realism', *Synthese*, Vol. 52.
- Durrant, R. G. (1970), 'Identity of Properties and the Definition of "Good"', *Australian Journal of Philosophy*, Vol. 48.
- Elliot, Robert (1987), 'Moral Realism and the Modal Argument', *Analysis*, Vol. 47.
- Falk, W. D. (1952 (1948)), "'Ought" and Motivation', *Readings in Ethical Theory*, ed. by Wilfrid Sellars and John Hospers, New York: Appleton-Century-Crofts.
- Feldman, Fred (1978), *Introductory Ethics*, Englewood Cliffs: Prentice-Hall.
- Finlay, Stephen (2004), 'The Conversational Practicality of Value Judgement', *The Journal of Ethics*, Vol. 8.
- Firth, Roderick (1952), 'Ethical Absolutism and the Ideal Observer',

- Philosophy and Phenomenological Research*, Vol. XII.
- Foot, Philippa (1978), *Virtue and Vice*, Los Angeles: University of California Press.
- Foot, Philippa (1978 (1972)), 'Morality as a System of Hypothetical Imperatives', *Virtue and Vice*, Los Angeles: University of California Press.
- Frankena, William K. (1976), *Perspectives on Morality. Essays by William K. Frankena*, ed. by K. E. Goodpaster, Notre Dame: University of Notre Dame Press.
- Frankena, William K. (1976 (1939)), 'The Naturalistic Fallacy', *Perspectives on Morality. Essays by William K. Frankena*, ed. by K. E. Goodpaster, Notre Dame: University of Notre Dame Press.
- Frankena, William K. (1976 (1958)), 'Obligation and Motivation in Recent Moral Philosophy', *Perspectives on Morality. Essays by William K. Frankena*, ed. by K. E. Goodpaster, Notre Dame: University of Notre Dame Press.
- Fumerton, Richard A. (1983), 'The Paradox of Analysis', *Philosophy and Phenomenological Research*, Vol. XLIII.
- Fumerton, Richard A. (1990), *Reason and Morality*, Ithaca: Cornell University Press.
- Gampel, Eric H. (1996), 'A Defence of the Autonomy of Ethics: Why Value Is Not Like Water', *Canadian Journal of Philosophy*, Vol. 26.
- Gampel, Eric H. (1997), 'Ethics, Reference, and Natural Kinds', *Philosophical Papers*, Vol. XXVI.
- Garcia, J. L. A. (1988), 'Relativism and Moral Divergence', *Metaphilosophy*, Vol. 19.
- Gibbard, Allan (1990), *Wise Choices, Apt Feelings*, Oxford: Clarendon Press.
- Goldman, Alan H. (1988), *Moral Knowledge*, London: Routledge.
- Grayling, A. C. (1987), 'Realism', *Cogito*, Vol. 4.
- Grice, Paul (1989), *Studies in the Ways of Words*, Cambridge, Mass.: Harvard University Press.
- Grice, Paul (1989a (1975)), 'Logic and Conversation', *Studies in the Ways of Words*, Cambridge, Mass.: Harvard University Press.
- Grice, Paul (1989b (1975)), 'Further Notes on Logic and Conversation', *Studies in the Ways of Words*, Cambridge, Mass.: Harvard University Press.
- Griffin, James (1992), 'Values: Reduction, Supervenience, and Explanation by Ascent', *Reduction, Explanation, and Realism*, ed. by David Charles and Kathleen Lennon, Oxford: Clarendon Press.
- Griffin, James (1996), *Value Judgement*, Oxford: Oxford University Press.
- Grimes, Thomas R. (1988), 'The Myth of Supervenience', *Pacific Philosophical Quarterly*, Vol. 69.
- Grimes, Thomas R. (1991), 'Supervenience, Determination and

- Dependency', *Philosophical Studies*, Vol. 62.
- Hare, R. M. (1952), *The Language of Morals*, Oxford: Clarendon Press.
- Hare, R. M. (1981), *Moral Thinking*, Oxford: Clarendon Press.
- Hare, R. M. (1997), *Sorting Out Ethics*, Oxford: Oxford University Press.
- Hare, R. M. (1989), *Essays in Ethical Theory*, Oxford: Clarendon Press.
- Hare, R. M. (1989 (1984)), 'Supervenience', *Essays in Ethical Theory*, Oxford: Clarendon Press.
- Hare, R. M. (1999), *Objective Prescriptions*, Oxford: Clarendon Press.
- Hare, R. M. (1999 (1993)), 'Objective Prescriptions', *Objective Prescriptions*, Oxford: Clarendon Press.
- Hare, R. M. (1999 (1996)), 'A New Kind of Ethical Naturalism?', *Objective Prescriptions*, Oxford: Clarendon Press.
- Harman, Gilbert (1975), 'Moral Relativism Defended', *Philosophical Review*, Vol. 84.
- Harman, Gilbert (1977), *The Nature of Morality*, Oxford: Oxford University Press.
- Harman, Gilbert (1984), 'Is there a Single True Reality?', *Morality, Reason and Truth. New Essays on the Foundations of Ethics*, ed. by David Copp and David Zimmerman, Totowa: Rowman & Allanheld.
- Harman, Gilbert (1986a), 'Moral Explanations of Natural Facts—Can Moral Claims Be Tested Against Moral Reality?', *The Southern Journal of Philosophy*, Vol. XXIV, Supplement: Spindel Conference 1986: Moral Realism, ed. by Norman Gillespie.
- Harman, Gilbert (1986b), 'Moral Agent and Impartial Spectator', *Lindley Lecture*, University of Kansas.
- Harman, Gilbert and Judith Jarvis Thomson (1996), *Moral Relativism and Moral Objectivity*, Oxford: Blackwell.
- Hatzimoysis, Anthony (2002), 'Analytical Descriptivism Revisited', *Ratio*, Vol. XV.
- Heil, John (1999), 'Multiple Realizability', *American Philosophical Quarterly*, Vol. 36.
- Holland, Sean (2001), 'Dispositional Theories of Value Meet Moral Twin Earth', *American Philosophical Quarterly*, Volume 38.
- Hooker, Brad (2000), 'Moral Particularism: Wrong and Bad', *Moral Particularism*, ed. by Brad Hooker and Margaret Olivia Little, Oxford: Oxford University Press.
- Hooker, C. A. (1981), 'Towards a General Theory of Reduction. Part II: Identity and Reduction', *Dialogue*, Vol. XX.
- Horgan, Terence (1993), 'From Supervenience to Superdupervenience: Meeting the Demands of a Material World', *Mind*, Vol. 102.
- Horgan, Terence and Mark Timmons (1991), 'New Wave Moral Realism Meets Moral Twin Earth', *Journal of Philosophical Research*, Vol. XVI.
- Horgan, Terence and Mark Timmons (1992a), 'Troubles on Moral Twin



- Earth: Moral Queerness Revived', *Synthese*, Vol. 92.
- Horgan, Terence and Mark Timmons (1992b), 'Troubles for New Wave Moral Semantic: The "Open Question Argument" Revived', *Philosophical Papers*, Vol. XXI.
- Horgan, Terence and Mark Timmons (1996), 'From Moral Realism to Moral Relativism in one Easy Step', *Critica*, Vol. XXVIII.
- Horgan, Terence and Mark Timmons (2000), 'Copping out on Moral Twin Earth', *Synthese*, Vol. 124.
- Horwich, Paul (1982), 'Three Forms of Realism', *Synthese*, Vol. 51.
- Horwich, Paul (1996), 'Realism and Truth', *Philosophical Perspectives*, 10, Metaphysics, ed. by James E. Tomberlin, Cambridge, M.A.: Blackwell.
- Hume, David (1978 (1888)), *A Treatise of Human Nature*, ed. by L. A. Selby-Bigge, Oxford: Clarendon Press.
- Jackson, Frank (1998), *From Metaphysics to Ethics*, Oxford: Oxford University Press.
- Jackson, Frank (2003) 'Cognitivism, A Priori Deduction and Moore', *Ethics*, Vol. 113.
- Jackson, Frank and Philip Pettit (1995), 'Moral Functionalism and Moral Motivation', *The Philosophical Quarterly*, Vol. 45.
- Johnston, Mark (1989), 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society*, Supplementary Volume LXIII.
- Kihlbom, Ulrik (2002), *Ethical Particularism. An Essay on Moral Reasons*, Stockholm: Almqvist & Wiksell.
- Kim, Jaegwon (1983), 'Supervenience and Supervenient Causation', *The Southern Journal of Philosophy*, Vol. XXII, Supplement: Spindel Conference 1983: Supervenience, ed. by Terence Horgan.
- Kim, Jaegwon (1993), *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Kim, Jaegwon (1993 (1984)), 'Concepts of Supervenience', *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Kim, Jaegwon (1993 (1990)), 'Supervenience as a Philosophical Concept', *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Kim, Jaegwon (1993 (1992)), 'Multiple Realization and the Metaphysics of Reduction', *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Kim, Jaegwon (1993 (1993)), 'The Nonreductivist's Troubles with Mental Causation', *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Kim, Jaegwon (1997), 'Moral Kinds and Natural Kinds: What's the Difference—for a Naturalist', *Philosophical Issues*, Vol. 8, Truth, ed. by Enrique Villanueva.
- King, Jeffrey C. (1990), 'What Is a Philosophical Analysis?', *Philosophical*

- Studies*, Vol. 90.
- Klagge, James (1984), 'An Alleged Difficulty Concerning Moral Properties', *Mind*, Vol. XCIII.
- Klagge, James (1987), 'Supervenience: Perspectives v. Possible Worlds', *The Philosophical Quarterly*, Vol. 37.
- Klagge, James (1988), 'Supervenience: Ontological and Ascriptive', *Australasian Journal of Philosophy*, Vol. 66.
- Korsgaard, Christine M. (1996), *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press.
- Korsgaard, Christine M. (1996 (1986)), 'Skepticism About Practical Reason', *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press.
- Korsgaard, Christine M. (1996 (1989)), 'Kant's Analysis of Obligation: The Argument of *Groundwork I*', *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press.
- Kraemer, Eric Russert (1991), 'On the Moral Twin Earth Challenge to New-Wave Moral Realism', *Journal of Philosophical Research*, Vol. XVI.
- Kripke, Saul (1980), *Naming and Necessity*, Oxford: Blackwell.
- Leiter, Brian (2001), 'Moral Facts and Best Explanations', *Social Philosophy and Policy*, Vol. 8.
- Lewis, David (1989), 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society*, Supplementary Volume LXIII.
- Lewy, C. (1965), 'G. E. Moore on the Naturalistic Fallacy', *Proceedings of the British Academy*, Vol. L.
- Lillehammer, Hallvard (1997), 'Smith on Moral Fetishism', *Analysis*, Vol. 57.
- Lillehammer, Hallvard (1999), 'Analytic Dispositionalism and Practical Reason', *Ethical Theory and Moral Practice*, Vol. 2.
- Lillehammer, Hallvard (2002), 'Moral Cognitivism', *Philosophical Papers*, Vol. 31.
- Lippert-Rasmussen, Kasper (1999), 'On Denying a Significant Version of the Constancy Assumption', *Theoria*, Vol. LXV.
- Lockie, Robert (1998), 'What's Wrong with Moral Internalism', *Ratio*, Vol. XI.
- Loeb, D. (1998), 'Moral Realism and the Argument from Disagreement', *Philosophical Studies*, Vol. 90.
- Lycan, William G. (1986), 'Moral Facts and Moral Knowledge', *The Southern Journal of Philosophy*, Vol. XXIV, Supplement: Spindel Conference 1986: Moral Realism, ed. by Norman Gillespie.
- Lyons, David (1982), 'Ethical Relativism and the Problem of Incoherence', *Relativism. Cognitive and Moral*, ed. by Michael Krausz and Jack Meiland, Notre Dame: University of Notre Dame Press.

- Macdonald, Cynthia (1989), *Mind-Body Identity Theories*, London: Routledge.
- Mackie, J. L. (1977), *Ethics. Inventing Right and Wrong*, London: Penguin Books.
- Mackie, J. L. (1980), *Hume's Moral Theory*, London: Routledge.
- Maddy, Penelope (1990), *Realism in Mathematics*, Oxford: Clarendon Press.
- Magnell, Thomas (1988), 'Moore's Attack on Naturalism', *Inquiries into Values*, ed. by Sander Lee.
- Majors, B. (2003), 'Moral Explanation and the Special Sciences', *Philosophical Studies*, Vol. 113.
- Marras, Ausonio (1993), 'Psychophysical Supervenience and Nonreductive Materialism', *Synthese*, Vol. 95.
- Martin, C. B. (1997), 'On the Need for Properties: The Road to Pythagoreanism and Back', *Synthese*, Vol. 112.
- McDowell, John (1978), 'Are Moral Requirements Hypothetical Imperatives?', *The Aristotelian Society*, Supplementary Volume, LII.
- McDowell, John (1979), 'Virtue and Reason', *Monist*, Vol. 62.
- McFetridge, I. G. (1985), 'Supervenience, Realism, Necessity', *The Philosophical Quarterly*, Vol. 35.
- McLaughlin, Brian P. (1995), 'Varieties of Supervenience', *Supervenience. New Essays*, ed. by Elias E. Savellos and Ümit D. Yalçın, Cambridge: Cambridge University Press.
- McNaughton, David (1988), *Moral Vision*, Oxford: Blackwell.
- McNaughton, David and Piers Rawling (2003), 'Can Scanlon Avoid Redundancy by Passing the Buck?', *Analysis*, Vol. 63.
- Merli, David (2002), 'Return to Moral Twin Earth', *Canadian Journal of Philosophy*, Vol. 32.
- Miller, Alexander (2003), *An Introduction to Contemporary Metaethics*, Cambridge: Polity Press.
- Miller, Richard W. (1985), 'Ways of Moral Learning', *The Philosophical Review*, Vol. XCIV.
- Milo, Ronald D. (1981), 'Moral Indifference', *The Monist*, Vol. 64.
- Moore, G. E. (1993 (1903)), *Principia Ethica*, with the preface to the second edition and other papers, edited with an introduction by Thomas Baldwin, Cambridge: Cambridge University Press.
- Moore, G. E. (1993 (1922)), 'The Conception of Intrinsic Value', in the volume with *Principia Ethica*, Cambridge: Cambridge University Press.
- Moore, G. E. (1942), 'A Reply to my Critics', *The Philosophy of G. E. Moore*, ed. by Paul Arthur Schilpp, Chicago: Northwestern University.
- Mumford, Stephen (1998), *Dispositions*, Oxford: Oxford University Press.
- Myers, C. Mason (1971), 'Moore's Paradox of Analysis', *Metaphilosophy*,

Vol. 2.

- Nagel, Thomas (1970), *The Possibility of Altruism*, New York: Oxford University Press.
- Nakhnikian, George (1963), 'On the Naturalistic Fallacy', *Morality and the Language of Conduct*, ed. by Hector-Neri Castañeda and George Nakhnikian, Detroit: Wayne State University Press.
- Nielsen, Kai (1974), 'Covert and Overt Synonymity: Brandt and Moore and the "Naturalistic Fallacy"', *Philosophical Studies*, Vol. 25.
- Nielsen, Kai (1984), 'Why Should I Be Moral? Revisited', *American Philosophical Quarterly*, Vol. 21.
- Noonan, H. W. (1987), 'Supervenience', *The Philosophical Quarterly*, Vol. 37.
- Nowell-Smith, P. H. (1954) *Ethics*, London: Penguin.
- Olson, Jonas (2002), 'Are Desires *De Dicto* Fetishistic?', *Inquiry*, Vol. 45.
- Owens, David (1989), 'Disjunctive Laws?', *Analysis*, Vol. 49.
- Parfit, Derek (1997), 'Reasons and Motivation', *The Aristotelian Society*, Supplementary Volume LXXI.
- Parfit, Derek (2003), 'Justifiability to Each Person', *Ratio*, Vol. XVI.
- Penczek, Alan (1997), 'Disjunctive Properties and Causal Efficacy', *Philosophical Studies*, Vol. 86.
- Pereboom, Derk and Hilary Kornblith (1991), 'The Metaphysics of Irreducibility', *Philosophical Studies*, Vol. 63.
- Persson, Ingmar (1997), 'Hume—Not a "Humean" About Motivation', *History of Philosophy Quarterly*, Vol. 14.
- Pettit, Philip (2000), 'Two Construals of Scanlon's Contractualism', *Journal of Philosophy*, Vol. XCVII.
- Pettit, Philip (2001), 'Embracing Objectivity in Ethics', *Objectivity in Law and Morals*, Cambridge: Cambridge University Press.
- Pettit, Philip, (2002), *Rules, Reasons and Norms*, Oxford: Oxford University Press.
- Pettit, Philip and Michael Smith (1990), 'Backgrounding Desire', *The Philosophical Review*, Vol. XCIX.
- Pigden, Charles R. (1991), 'Naturalism', *A Companion to Ethics*, ed. by Peter Singer, Oxford: Blackwell.
- Platts, Mark de Bretton (1979), *Ways of Meaning*, London: Routledge.
- Pressler, Jonathan (1988), 'Is There any Basis for Moral Scepticism?', *Australasian Journal of Philosophy*, Vol. 66.
- Prichard, H. A. (1952 (1912)), 'Does Moral Philosophy Rest on a Mistake?', *Readings in Ethical Theory*, ed. by Wilfrid Sellars and John Hospers, New York: Appleton-Century-Crofts.
- Putnam, Hilary (1975), *Mind, Language and Reality. Philosophical Papers*, Volume 2, Cambridge: Cambridge University Press.
- Putnam, Hilary (1975a), 'Language and Philosophy', *Mind, Language and*

- Reality. Philosophical Papers*, Volume 2, Cambridge: Cambridge University Press.
- Putnam, Hilary (1975b), 'The Meaning of "Meaning"', *Mind, Language and Reality. Philosophical Papers*, Volume 2, Cambridge: Cambridge University Press.
- Putnam, Hilary (1981), *Reason, Truth and History*, Cambridge: Cambridge University Press.
- Quinn, Warren S. (1978), 'Moral and Other Realisms: Some Initial Difficulties', *Vales and Morals*, ed. by A. I. Goldman and J. Kim, Dordrecht: D. Reidel.
- Quinn, Warren S. (1986), 'Truth and Explanation in Ethics', *Ethics*, Vol. 96.
- Rabinowicz, Wlodek and Toni Rønnow-Rasmussen (2000), 'A Distinction in Value: Intrinsic and For Its Own Sake', *Proceedings of the Aristotelian Society*, Vol. 100.
- Railton, Peter (1989), 'Naturalism and Prescriptivity', *Social Philosophy and Policy*, Vol. 7.
- Railton, Peter (1993a), 'What the Non-Cognitivist Helps Us to See the Naturalist Must Help Us to Explain', *Reality, Representation, and Projection*, ed. by John Haldane and Crispin Wright, Oxford: Oxford University Press.
- Railton, Peter (1993b), 'Reply to David Wiggins', *Reality, Representation, and Projection*, ed. by John Haldane and Crispin Wright, Oxford: Oxford University Press.
- Railton, Peter (1996a), 'Moral Realism: Prospects and Problems', *Moral Knowledge? New Readings in Moral Epistemology*, ed. by Walter Sinnott-Armstrong and Mark Timmons, Oxford: Oxford University Press.
- Railton, Peter (1996b), 'Subject-ive and Objective'. *Truth in Ethics*, ed. by Brad Hooker, Oxford: Blackwell.
- Railton, Peter (1998), 'Moral Explanation and Moral Objectivity', *Philosophy and Phenomenological Research*, Vol. LVIII.
- Raz, Joseph (1975), *Practical Reason and Norms*, London: Hutchinson of London.
- Raz, Joseph (2000), 'The Truth in Particularism', *Moral Particularism*, ed. by Brad Hooker and Margaret Olivia Little, Oxford: Oxford University Press.
- Regan, Tom (1972), 'Moore's Accounts of "Right"', *Dialogue*, Vol. XI.
- Rønnow-Rasmussen, Toni (1993), *Logic, Facts and Representation. An Examination of R. M. Hare's Moral Philosophy*, Lund: Lund University Press.
- Roojen, Mark van (1996), 'Moral Functionalism and Moral Reductionism', *The Philosophical Quarterly*, Vol. 46.

- Ross, W. D. (1939), *Foundations of Ethics*, Oxford: Clarendon Press.
- Ross, W. D. (1988 (1930)), *The Good and the Right*, Oxford: Indianapolis: Hackett.
- Ryle, Gilbert (1966 (1949)), *The Concept of Mind*, Harmondsworth: Penguin.
- Sadler, Brook J. (2003), 'The Possibility of Amoralism: A Defence Against Internalism', *Philosophy*, Vol. 78.
- Sayre-McCord, Geoffrey (1988 (1986)), 'The Many Moral Realisms', *Essays on Moral Realism*, ed. by Geoffrey Sayre-McCord, Ithaca: Cornell University Press.
- Sayre-McCord, Geoffrey (1988), 'Moral Theory and Explanatory Impotence', *Essays on Moral Realism*, ed. by Geoffrey Sayre-McCord, Ithaca: Cornell University Press.
- Sayre-McCord, Geoffrey (1991), 'Being a Realist About Relativism (in Ethics)', *Philosophical Studies*, Vol. 61.
- Sayre-McCord, Geoffrey (1992), 'Normative Explanations', *Philosophical Perspective*, 6, Ethics.
- Sayre-McCord, Geoffrey (1996), 'Coherentist Epistemology and Moral Theory', *Moral Knowledge? New Readings in Moral Epistemology*, ed. by Walter Sinnott-Armstrong and Mark Timmons, Oxford: Oxford University Press.
- Sayre-McCord, Geoffrey (1997a), "'Good" on Twin Earth', *Philosophical Issues*, Vol. 8, Truth, ed. by Enrique Villanueva.
- Sayre-McCord, Geoffrey (1997b), 'Replies to Comments', *Philosophical Issues*, Vol. 8, Truth, ed. by Enrique Villanueva.
- Scanlon, T. M. (1998), *What We Owe to Each Other*, Cambridge, Mass.: Harvard University Press.
- Schiffer, Stephen (1990), 'Meaning and Value', *The Journal of Philosophy*, Vol. LXXXVII.
- Schurz, Gerhard (1997), *The Is-Ought Problem*, Dordrecht: Kluwer Academic Publishers.
- Schwarz, Stephen D. (1971), 'Does Prichard's Essay Rest on a Mistake?', *Ethics*, Vol. 81.
- Seager, William (1991), 'Disjunctive Laws and Supervenience', *Analysis*, Vol. 51.
- Searle, John (1969), *Speech Acts. An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.
- Searle, John (1979), *Expression and Meaning. Studies in the Theory of Speech Acts*, Cambridge: Cambridge University Press.
- Searle, John (1979 (1975)), 'Indirect Speech Acts', *Expression and Meaning. Studies in the Theory of Speech Acts*, Cambridge: Cambridge University Press.
- Shafer-Landau, Russ (1995), 'Vagueness, Borderline Cases and Moral Realism', *American Philosophical Quarterly*, Vol. 32.

- Shafer-Landau, Russ (1998), 'Moral Judgment and Moral Motivation', *The Philosophical Quarterly*, Vol. 48.
- Shafer-Landau, Russ (2003), *Moral Realism. A Defence*, Oxford: Clarendon Press.
- Sidelle, Alan (1989), *Necessity, Essence and Individuation*, Ithaca: Cornell University Press.
- Sidgwick, Henry (1981 (1907)), *The Methods of Ethics*, Cambridge, Mass.: Hackett Publishing Company.
- Sinnott-Armstrong, Walter (1984), "'Ought" Conversationally Implies "Can"', *The Philosophical Review*, Vol. XCIII.
- Sinnott-Armstrong, Walter (1996), 'Moral Skepticism and Justification', *Moral Knowledge? New Readings in Moral Epistemology*, ed. by Walter Sinnott-Armstrong and Mark Timmons, Oxford: Oxford University Press.
- Sinnott-Armstrong, Walter (1999), 'Some Varieties of Particularism', *Metaphilosophy*, Vol. LLC.
- Smith, M. B. E. (1972), 'Indifference and Moral Acceptance', *American Philosophical Quarterly*, Vol. 9.
- Smith, Michael (1991), 'Realism', *A Companion to Ethics*, ed. by Peter Singer, Oxford: Blackwell.
- Smith, Michael (1994), *The Moral Problem*, Oxford: Blackwell.
- Smith, Michael (1996), 'The Argument for Internalism: Reply to Miller', *Analysis*, Vol. 56.
- Smith, Michael (1997), 'In Defence of *The Moral Problem*: A Reply to Brink, Copp and Sayre-McCord', *Ethics*, Vol. 108.
- Smith, Michael (1999), 'Does the Evaluative Supervene on the Natural?', *Well-Being and Morality. Essays in Honour of James Griffin*, ed. by Roger Crisp and Brad Hooker, Oxford: Clarendon Press.
- Smith, Michael (2001a), 'The Resentment Argument', *Exploring Practical Philosophy. From Action to Values*, ed. by Jonas Josefsson, Björn Petersson and Toni Rønnow-Rasmussen, Hampshire: Ashgate.
- Smith, Michael (2001b), 'Some Not-much-discussed Problems for Non-cognitivism in Ethics', *Ratio*, Vol. XIV.
- Snare, Frank (1975), 'The Open Question as a Linguistic Test', *Ratio*, Vol. 17.
- Sober, Elliott (1982a), 'Realism and Independence', *Noûs*, Vol. XVI.
- Sober, Elliott (1982b), 'Why Logically Equivalent Predicates May Pick out Different Properties', *American Philosophical Quarterly*, Vol. 19.
- Sosa, David (2004), 'A Big, Good Thing', *Noûs*, Vol. 38.
- Sosa, Ernest (1997), 'Water, Drink, and "Moral Kinds"', *Philosophical Issues*, Vol. 8, Truth, ed. by Enrique Villanueva.
- Stevenson, Charles L. (1944), *Ethics and Language*, New Haven: Yale University Press.

- Stevenson, Charles L. (1963), *Facts and Value*, New Haven: Yale University Press.
- Stevenson, Charles L. (1963 (1937)), 'The Emotive Meaning of Ethical Terms', *Facts and Value*, New Haven: Yale University Press.
- Stevenson, Charles L. (1963 (1938)), 'Persuasive Definitions', *Facts and Value*, New Haven: Yale University Press.
- Stevenson, Charles L. (1963 (1948)), 'The Nature of Ethical Disagreement', *Facts and Value*, New Haven: Yale University Press.
- Stocker, Michael (1979), 'Desiring the Bad: An Essay in Moral Psychology', *The Journal of Philosophy*, Vol. LXXVI.
- Strandberg, Caj (1999), 'Dispositional Moral Properties and Moral Motivation', *Theoria*, Vol. LXV.
- Strandberg, Caj (2000), 'Aristotle's Internalism in the *Nicomachean Ethics*', *The Journal of Value Inquiry*, Vol. 34.
- Strandberg, Caj (2004), 'In Defence of the Open Question Argument', *The Journal of Ethics*, Vol. 8.
- Stratton-Lake, Philip (1999), 'Why Externalism Is Not a Problem for Ethical Intuitionists', *Proceedings of the Aristotelian Society*, Vol. XCIX.
- Stratton-Lake, Philip (2000), *Kant, Duty and Moral Worth*, London: Routledge.
- Stratton-Lake, Philip (2003), 'Scanlon's Contractualism and the Redundancy Objection', *Analysis*, Vol. 63.
- Sturgeon, Nicholas L. (1984), 'Moral Explanations', *Morality, Reason and Truth. New Essays on the Foundations of Ethics*, ed. by David Copp and David Zimmerman, Totowa: Rowman & Allanheld.
- Sturgeon, Nicholas L. (1986a), 'Harman on Moral Explanations of Natural Facts', *The Southern Journal of Philosophy*, Vol. XXIV, Supplement: Spindel Conference 1986: Moral Realism, ed. by Norman Gillespie.
- Sturgeon, Nicholas L. (1986b), 'What Difference Does It Make Whether Moral Realism Is True?', *The Southern Journal of Philosophy*, Vol. XXIV, Supplement: Spindel Conference 1986: Moral Realism, ed. by Norman Gillespie.
- Sturgeon, Nicholas L. (1992), 'Nonmoral Explanations', *Philosophical Perspective*, 6, Ethics.
- Sturgeon, Nicholas L. (1994), 'Moral Disagreement and Moral Relativism', *Social Philosophy and Policy*, Vol. 11.
- Sturgeon, Nicholas L. (2002), 'Ethical Intuitionism, Ethical Naturalism', *Ethical Intuitionism*, ed. by Philip Stratton-Lake, Oxford: Clarendon Press.
- Sturgeon, Nicholas L. (2003), 'Moore on Ethical Naturalism', *Ethics*, Vol. 113.
- Svavarsdóttir, Sigrún (1999), 'Moral Cognitivism and Motivation', *The Philosophical Review*, Vol. 108.



- Tännsjö, Torbjörn (1990), *Moral Realism*, Savage: Rowman and Littlefield.
- Teller, Paul (1983), 'Comments on Kim's Paper', *The Southern Journal of Philosophy*, Vol. XXII, Supplement: Spindel Conference 1983: Supervenience, ed. by Terence Horgan.
- Tenenbaum, Sergio (2000), 'Ethical Internalism and Glaucon's Question', *Noûs*, Vol. 34.
- Tersman, Folke (2002), *The Metaethical Significance of Moral Disagreement*, manuscript.
- Timmons, Mark (1990), 'On the Epistemic Status of Considered Moral Judgments', *The Southern Journal of Philosophy*, Vol. XXIX, Supplement: Spindel Conference 1990: Moral Epistemology, ed. by Norman Gillespie.
- Timmons, Mark (1999), *Morality without Foundations*, Oxford: Oxford University Press.
- Van Cleve, James (1990), 'Supervenience and Closure', *Philosophical Studies*, Vol. 58.
- Warnock, G. J. (1967), *Contemporary Moral Philosophy*, London: MacMillan.
- Wedgwood, Ralph (1999), 'The Price of Non-Reductive Moral Realism', *Ethical Theory and Moral Practice*, Vol. 2.
- Wedgwood, Ralph (2001), 'Conceptual Role Semantics for Moral Terms', *The Philosophical Review*, Vol. 101.
- Wellman, Carl (1975), 'Ethical Disagreement and Objective Truth', *American Philosophical Quarterly*, Vol. 12.
- Werner, Richard (1983), 'Ethical Realism', *Ethics*, Vol. 93.
- Wiggins, David (1987), *Needs, Values, Truth. Essays in the Philosophy of Value*, Oxford: Blackwell.
- Wiggins, David (1987 (1983)), 'Truth as Predicated of Moral Judgments', *Needs, Values, Truth. Essays in the Philosophy of Value*, Oxford: Blackwell.
- Wiggins, David (1991), 'Moral Cognitivism, Moral Relativism and Motivating Moral Beliefs', *Proceedings of the Aristotelian Society*, Vol. 91.
- Williams, Bernard (1981), *Moral Luck*, Cambridge: Cambridge University Press.
- Williams, Bernard (1981 (1980)), 'Internal and External Reasons', *Moral Luck*, Cambridge: Cambridge University Press.
- Williams, Bernard (1985), *Ethics and the Limits of Philosophy*, Glasgow: Fontana Press.
- Wong, David (1984), *Moral Relativity*, Berkeley: University of California Press.
- Wreen, Michael (1985), 'Mackie on the Objectivity of Values', *Dialectica*, Vol. 39.
- Wright, Crispin (1992), *Truth and Objectivity*, Cambridge, Mass.: Harvard

- University Press.
- Wright, Crispin (1996), 'Truth in Ethics', *Truth in Ethics*, ed. by Brad Hooker, Oxford: Blackwell.
- Yasenchuk, Ken (1995), 'Assimilative Moral Realism and Supervenience', *Dialogue*, Vol. XXXIV.
- Zangwill, Nick (1995), 'Moral Supervenience', *Midwest Studies in Philosophy*, Volume XX, Moral Concepts, ed. by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein.
- Zangwill, Nick (1997), 'Explaining Supervenience: Moral and Mental', *Journal of Philosophical Research*, Vol. XXII.
- Zangwill, Nick (1998), 'Supervenience, Reduction, and Infinite Disjunction', *Philosophia*, Vol. 26.
- Zangwill, Nick (2000), 'Against Analytic Moral Functionalism', *Ratio*, XIII.
- Zangwill, Nick (2003), 'Externalist Moral Motivation', *American Philosophical Quarterly*, Vol. 40.
- Zimmerman, David (1984), 'Moral Realism and Explanatory Necessity', *Morality, Reason and Truth. New Essays on the Foundations of Ethics*, ed. by David Copp and David Zimmerman, Totowa: Rowman & Allanheld.
- Zimmerman, Michael (2001), *The Nature of Intrinsic Value*, Lanham: Rowman and Littlefield.