# **Moral Disagreement and Moral Indeterminacy**

Laura-Janett Vasile

Submitted to the Central European University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy Supervisor: Professor János Kis

Department of Philosophy Central European University, Budapest 2010 I hereby declare that the dissertation contains no materials accepted for any other degrees in any other institutions of higher education. Also, I declare that the dissertation contains no materials previously written and/or published by another person, except where appropriate acknowledgment is made in the form of bibliographical references.

Laura-Janett Vasile February 2010

## Acknowledgments

I would like to thank my supervisor, János Kis, for the intellectual guidance with which he provided me over the years, as well as for his steadfast support and patience. Thanks are also due to Andrei Apostol, István Aranyosi, Cristian Constantinescu, Anca Gheaus, Shlomit Harrosh, Mara Marin, Bogdan Popescu, Serban Popescu-Szekely, George Turturea, and Camil Ungureanu, for helpful discussions and criticisms.

## Abstract

My aim in this dissertation is to provide an answer to the question whether the argument from moral disagreement undermines the plausibility of moral objectivism. I focus on a recurrent argument in the recent metaethical literature, according to which the moral objectivist can respond to the challenge posed by the pervasiveness and intractability of moral disagreement by appealing to the idea of moral indeterminacy. A substantial part of the dissertation is devoted to showing that this response to the argument from moral disagreement is not free of difficulties. One of the main worries has to do with the fact that, since moral indeterminacy cannot be a pervasive phenomenon, it seems that the best strategy available to the objectivist is to maintain that there is also a considerable amount of convergence in moral belief, and that this convergence in belief is best explained on objectivist assumptions. Yet, if responding to the argument from moral disagreement involves defending both the claim that there are indeterminate moral issues and the claim that convergence in moral belief can suitably be regarded as the result of a reliable epistemic mechanism, a worry might arise as to whether an adequate defense of both the aforementioned claims can be mounted from the perspective of the same version of moral objectivism. The worry is that defending the former claim may be easier if one embraces a version of cognitivism which delivers a weaker form of objectivity, while defending the latter claim may be easier if a version of cognitivism which delivers a stronger form of objectivity is vindicated. I claim that moral realism has the resources to account for both the aforementioned claims. On the one hand, I argue that moral realism has no trouble accommodating moral indeterminacy. On the other hand, I argue that moral facts figure in the best explanation of our moral beliefs. Moreover, I argue that despite common assumptions to the contrary, the claim that moral facts can play causal explanatory roles can consistently be endorsed by both naturalists and nonnaturalists.

## Contents

Acknowledgments	i
Abstract	ii
Chapter 1	1
Introduction	1
1.1 Moral disagreement and moral indeterminacy	1
1.2 Overview	4
Chapter 2	7
The argument from moral disagreement	7
2.1 Moral objectivism	7
2.2 Noncognitivism and error theory	11
2.3 Objectivity and convergence	17
2.4 The argument from moral disagreement	
2.5 Moral phenomenology	
2.6 Realism and indeterminacy.	
2.6.1 Dummett on realism and indeterminacy	
2.6.2 Comparative indeterminacy	
2.6.3 Alethic indeterminacy.	
2.6.4 Response-dependence and indeterminacy	64
Chapter 3	
The argument from explanatory impotence	
3.1 Harman's argument	
3.2 Causal and explanatory reasoning	
3.3 Ethical naturalism	
3.4 Moral supervenience	
3.5 Vindicatory explanations	
	100
Chapter 4	
Chapter 4 Value incommensurability	
Chapter 4 Value incommensurability 4.1 Value incommensurability	
Chapter 4 Value incommensurability 4.1 Value incommensurability 4.2 What items can be incommensurable?	
Chapter 4 Value incommensurability 4.1 Value incommensurability 4.2 What items can be incommensurable? 4.3 The covering value requirement	<b>109</b> <b>109</b> 109 109 
Chapter 4 Value incommensurability 4.1 Value incommensurability 4.2 What items can be incommensurable? 4.3 The covering value requirement 4.4 Arguments for incommensurability	<b>109</b> <b>109</b> 109 110 110 118 120
Chapter 4 Value incommensurability 4.1 Value incommensurability 4.2 What items can be incommensurable? 4.3 The covering value requirement 4.4 Arguments for incommensurability 4.4.1 The argument from multiple rankings	109 109 109 110 110 118 120 120
Chapter 4	109 109 109 110 110 118 120 120 123
Chapter 4	109 109 109 110 118 120 120 123 125
Chapter 4	109 109 109 110 110 118 120 120 123 125 127
Chapter 4	109 109 109 110 118 120 120 123 125 125 127 130
Chapter 4	109 109 109 109 110 110 120 120 123 125 127 130 135
Chapter 4	109 109 109 109 110 110 120 120 123 125 127 130 135 140
Chapter 4	109 109 109 109 110 118 120 120 123 125 127 130 135 140 140
Chapter 4	109 109 109 109 110 118 120 120 123 125 127 130 135 140 140 140
Chapter 4	109       109       109       110       110       110       110       110       110       120       120       123       125       127       130       135       140       140       148
Chapter 4	109       109       109       109       110       110       110       110       110       110       110       120       120       123       125       127       130       135       140       140       148       151
Chapter 4	109       109       109       109       110       110       110       110       120       120       123       125       127       130       135       140       140       148       151       157
Chapter 4	109       109       109       109       110       118       120       120       123       125       127       130       135       140       140       148       151       157       161
Chapter 4	109       109       109       109       110       118       120       121       120       121       120       123       125       127       130       135       140       140       148       151       157       161
Chapter 4.     Value incommensurability     4.1 Value incommensurability.     4.2 What items can be incommensurable?     4.3 The covering value requirement     4.4 Arguments for incommensurability     4.4.1 The argument from multiple rankings     4.4.2 Multiple rankings and rough equality     4.4.3 The small-improvement argument     4.4.4 The epistemic objection     4.5 Parity.     4.6 Incommensurability as vagueness.     Chapter 5     Moral disagreement, value incommensurability and moral dilemmas     5.1 Ideal agents and moral indeterminacy.     5.2 Value incommensurability and moral dilemmas.     5.3 Defining moral dilemmas.     5.4 Value incommensurability, irresolvable moral conflicts and moral wrongdoing	109     109     109     109     110     118     120     123     125     127     130     135     140     140     140     141     151     157     161     164
Chapter 4.     Value incommensurability     4.1 Value incommensurability     4.2 What items can be incommensurable?     4.3 The covering value requirement     4.4 Arguments for incommensurability     4.4 Arguments for incommensurability     4.4.1 The argument from multiple rankings     4.4.2 Multiple rankings and rough equality     4.4.3 The small-improvement argument     4.4.4 The epistemic objection     4.5 Parity	109     109     109     109     110     118     120     123     125     127     130     135     140     140     140     141     140     141     151     157     161     164     164
Chapter 4	109     109     109     109     110     118     120     123     125     127     130     135     140     161     162     164     164     164     164
Chapter 4	109     109     109     109     110     118     120     123     125     127     130     135     140     161     162     163     164     164     164     166     169
Chapter 4	$\begin{array}{c} 109 \\ 109 \\ 109 \\ 109 \\ 109 \\ 110 \\ 118 \\ 120 \\ 123 \\ 125 \\ 127 \\ 130 \\ 135 \\ 140 \\ 140 \\ 148 \\ 151 \\ 157 \\ 161 \\ 157 \\ 161 \\ 164 \\ 164 \\ 166 \\ 169 \\ 173 \\ 173 \\ 109 \\ 100 \\$

## Chapter 1 Introduction

## 1.1 Moral disagreement and moral indeterminacy

How should we explain the fact that moral disagreements cannot always be settled to the satisfaction of all rational and well-informed persons?

Various accounts of the pervasiveness and intractability of moral disagreement are informed by two basic metaethical positions. On the one hand, different versions of cognitivism hold that the failure to achieve a moral consensus is insufficient evidence against the claim that there are genuine moral truths. On this view, whenever moral disagreement persists, at least one of the parties involved is mistaken. The implicit claim is that any moral controversy can be settled, once the parties have enough time to engage in the debate and no cognitive failure intervenes. On the other hand, noncognitivists insist that the extent and apparent intractability of moral disagreement are best explained by denying that moral judgments are the kind of statements which can have a truth-value. If our evaluative judgments were to mirror an independent moral reality, the argument goes, why not hope for a consensus similar to the one that we reach on empirical matters? Since this hope is constantly frustrated, we are led to the conclusion that morality is just a matter of projecting our own feelings onto a value-free reality.<sup>1</sup>

Admittedly, at least some cases of moral disagreement can be accounted for in cognitivist terms, i.e. in terms of the parties' cognitive shortcomings. However, by assigning cognitive shortcomings such a significant role in explaining disagreement, cognitivists seem to eschew the real issue confronting them. As Russ Shafer-Landau (1994: 332) rightly points out, since it seems that moral disagreement may persist even among *fully informed, rational* agents, the real challenge that the cognitivist has to meet is to account for hypothetical disagreement among ideally placed agents, rather than for actual, seemingly intractable disagreement. Hence, what the cognitivist has to carry out is the apparently paradoxical task of showing how ideally placed agents can disagree about some moral issue without any of them being mistaken.

A recurrent argument in the recent metaethical literature purports to show that the difficulties that cognitivists have in explaining disagreement among fully informed, rational agents can be alleviated. According to this argument, morality fails to yield a unique and determinate answer to every practical question. Once we accept that there are indeterminate moral issues, there is no need to explain disagreement in terms of the parties' cognitive shortcomings. Instead, moral disagreements can be understood as cases in which morally relevant considerations do not add up to a uniquely correct answer about what ought to be done.

My aim in this dissertation is to provide an assessment of the merits of the cognitivist's reply to the argument from moral disagreement. More specifically, I focus on the question whether the appeal to the idea of moral indeterminacy can assist the

<sup>&</sup>lt;sup>1</sup> As it will become apparent in the following chapter, the argument from moral disagreement is also relied upon by error theorists and relativists to support their views. However, for the sake of clarity, I only refer here to the opposition between cognitivism and noncognitivism.

cognitivist in dealing with the challenge posed by the pervasiveness and intractability of moral disagreement.

Although many authors suggest that, given the possibility that some moral issues are indeterminate, intractable moral disagreement poses no threat to cognitivism, only few of them deal with this idea in a detailed manner.<sup>2</sup> However, this line of defense against the argument from disagreement leads to various difficulties. One such difficulty has to do with the fact that cognitivists typically claim that moral phenomenology lends support to a cognitivist account of disagreement. In particular, they point out that what lies in the background of persistent moral disagreement is a shared assumption about there being a right answer – otherwise the parties would not stick to their points of view and disagreement would vanish. Yet, this claim does not sit well with the further claim that disagreement occurs when a moral question has no determinate answer. So, the cognitivist owes us at least an explanation of how to reconcile the claim that cognitivism does justice to the phenomenology of moral thought with the claim that disagreement arises out of indeterminacy. Another, related difficulty is that agents who suffer from no cognitive shortcoming would have to be aware of all morally relevant facts, including those which can render a situation morally indeterminate. Thus, the cognitivist has to explain why such agents would continue to disagree about how to settle a moral issue instead of just agreeing on its indeterminate character.<sup>3</sup>

Still another difficulty may stem from the fact that an adequate response to the argument from moral disagreement would involve defending not just (a) the claim that some moral issues are indeterminate and that our failure to convergence on a moral verdict is sometimes due to such indeterminacy, but also (b) the claim that there is a considerable amount of convergence in moral belief and that this convergence is best explained on cognitivist assumptions. As it will become clear in Chapter 2, the cognitivist is bound to accept both that indeterminacy cannot be a pervasive phenomenon and that making extensive use of the idea that our cognitive shortcomings prevent us from reaching a consensus on moral matters would undermine his position. Therefore, the best strategy available to the cognitivist is to stress that there is also a considerable amount of convergence in moral belief. Yet, one can argue that such convergence can be explained without assuming that there are genuine moral truths (for instance, it could be explained by reference to our shared upbringing). In order to put this worry to rest, the cognitivist has to show that convergence in moral belief can suitably be regarded as the result of a reliable epistemic mechanism. Thus, whether the cognitivist can successfully respond to the argument from disagreement depends in part on whether he can prove that moral beliefs are produced by a reliable epistemic mechanism.

The point is not just that the question whether cognitivism has the resources to respond to the challenge posed by moral disagreement cannot be treated independently of other metaethical questions. If responding to the argument from disagreement involves defending both (a) and (b), a worry might arise as to whether an adequate defense of both the aforementioned claims can be mounted from the perspective of the same version of cognitivism. The worry is that cognitivism comes in many different versions, and that defending (a) may be easier if one embraces a version of cognitivism which delivers a weaker form of objectivity, while, on the other hand, defending (b) may be easier if a

<sup>&</sup>lt;sup>2</sup> A noteworthy exception is Shafer-Landau (1994).

<sup>&</sup>lt;sup>3</sup> For emphasizing this difficulty, see Shafer-Landau (1994).

version of cognitivism which delivers a stronger form of objectivity is vindicated. For instance, one can hold that, in order to prove that moral beliefs are the result of a reliable belief-producing mechanism, the cognitivist has to put forward an argument to the effect that moral beliefs are produced by a causal mechanism. Furthermore, one can point out that only naturalistic moral realism can accommodate such a picture.

Yet, it is far from obvious that this version of cognitivism can also accommodate the idea that disagreement occurs when a moral question has no determinate answer. One can argue, for instance, that constructivist versions of cognitivism are better suited to accommodate this idea. According to a suggestion that I will consider in Chapter 5, constructivism is uniquely well placed to explain why fully informed, rational agents would disagree about how to settle a moral issue instead of just agreeing on its indeterminate character. Given that, on a constructivist view, there are no moral truths prior to or independent of the deliverances of some idealized agents, there is no way of telling whether a situation is morally indeterminate independently of the deliverances of such agents. In other words, there is no reason to suppose that idealized agents are aware of the indeterminate character of any particular situation. Thus, constructivism would have no trouble explaining why idealized agents are not privy to information of the sort that would undermine the very possibility of disagreement. The upshot of all these considerations is that unless the cognitivist offers a thorough defense of both (a) and (b), he cannot allay the suspicion that his argument tacitly trades on intuitions that support different versions of cognitivism. Yet surely, in order to prove that cognitivism has the resources to respond to the challenge posed by moral disagreement, one would have to show that there is at least one version of cognitivism that has the resources to account for both (a) and (b).

My interest in the argument from moral disagreement has been spurred by the fact that many authors suggest that this argument loses its bite once we concede that some moral issues are indeterminate. While this suggestion has become virtually commonplace in the metaethical literature<sup>4</sup>, the difficulties that beset any attempt to prove that moral disagreement is due to indeterminacy are rarely addressed. However, this line of response to the argument from disagreement will remain unconvincing unless these difficulties are thoroughly addressed.

Moreover, the appeal to the idea of moral indeterminacy recasts the debate between cognitivists and noncognitivists in new terms. Most significantly, once this idea is in play, the burden of proof seems to fall upon the cognitivist, on at least two counts. First, the cognitivist must offer an account of the sources of moral indeterminacy. Second, he has to prove that some of the most compelling cases of practical conflict can be translated in his terms. Although it can be argued that these two tasks do not make the proper object of metaethics, I believe that, unless it is seconded by an encompassing attempt to deal with these two tasks, the cognitivist proposal does not do much work on its own.

That metaethics witnesses an increasing interest in the notion of moral indeterminacy is at least partly explained by the fact that several other, related notions are

<sup>&</sup>lt;sup>4</sup> For this suggestion, see Brink (1984: 116–7; 1989: 202), Boyd (1988: 199–201; 212–4), Wiggins (1990– 91), Wolf (1992), Schafer-Landau (1994; 2003: ch. 9), Sturgeon (1994), Thomson (1996: 154, 205–6), Sosa (2001), Gert (2005: 4, 311), and Thomas (2006: 252).

at the forefront of current philosophical debate. Value incommensurability has been a focus of recent philosophical interest, and to the extent that we find plausible the idea that some moral issues are indeterminate, this is also due to there being an extensive body of literature aimed at showing that the value of certain options cannot be ranked against each other as "better than" or "equally good". Also, the fact that over the last three decades numerous philosophers have attempted to prove that there are genuine moral dilemmas has certainly played an important role in our finding the idea of moral indeterminacy so familiar.

It is important to stress, however, that the view according to which there are cases when the traditional trichotomy of relations "better than", "worse than" and "equally good" fails is highly controversial. Moreover, we should keep in mind that the nature and significance of moral dilemmas has been the object of a heated philosophical debate. So, insofar as the plausibility of the cognitivist's response to the argument from disagreement depends on whether there are cases of comparative moral indeterminacy, the cognitivist cannot avoid engaging in such debates. This is important especially since, in contrast with noncognitivists who typically claim that metaethical issues are completely independent of ethical issues, cognitivists call into question this independence and argue that our answer to many metaethical questions turns on how we answer certain ethical questions.<sup>5</sup>

#### **1.2 Overview**

Chapter 2 begins by setting the stage for the evaluation of the argument from moral disagreement. A natural first step is to identify the target of the argument. I have already pointed out that noncognitivists contend that the distinctive nature of moral disagreement provides a secure basis for rejecting cognitivism. Yet, as it will become clear in Chapter 2, considerations having to do with the pervasiveness and apparent intractability of moral disagreement are also relied upon by error theorists and moral relativists to support their views. So, one can say that the argument from disagreement is targeted primarily at moral objectivism and that the point of contention between noncognitivists, error theorists and moral relativists is precisely which objectivist tenet has to be rejected. In light of these considerations, the question I began with, i.e. whether the appeal to the idea of moral disagreement, can be rephrased to ask whether the argument from moral disagreement undermines the plausibility of objectivist versions of cognitivism, and in particular, whether the appeal to the idea of moral indeterminacy can help the objectivist win this debate. The remainder of the dissertation is devoted to answering this question.

One of the aims of Chapter 2 is to establish what sorts of considerations can legitimately be invoked by the objectivist in replying to the argument from disagreement. In order to do this, I take up the question whether certain versions of cognitivism (such as constructivist theories or response-dependent theories) count as objectivist in the sense that concerns us when discussing the argument from disagreement. I argue that it would

<sup>&</sup>lt;sup>5</sup> For emphasizing that we should view metaethical issues as continuous with ethical issues, see Brink (1989: 4–5) and Smith (1994: 3, 202).

be a mistake to think that (nonrelativist) constructivism, as opposed to realism, cannot properly be counted as objectivist. Furthermore, I argue that it would be a mistake to think that, in responding to the argument from disagreement, one cannot legitimately invoke considerations that derive from a broadly anthropocentric view of the nature of morality.

A major part of Chapter 2 is devoted to rebutting two of the most important objections that the proponents of the view that lack of moral consensus is due to there being no determinate answer to the question under considerations have to deal with. First, one can argue that, if the realist contends that the argument from disagreement loses its bite once we concede that some moral issues are indeterminate, this leads to the following problem. While the realist claims that his metaethical view does justice to moral phenomenology, by arguing that disagreement occurs when a moral question has no determinate answer, he ends up suggesting just the opposite. So, the realist owes us an explanation of how to reconcile the claim that his metaethical view accommodates the phenomenology of moral disagreement with the claim that disagreement arises out of indeterminacy. I argue that moral realism has the resources to answer this objection. The second worry about the view that lack of consensus on some moral issues is due to indeterminacy is that realism and indeterminacy are incompatible. In Chapter 2, I distinguish between alethic indeterminacy and comparative indeterminacy, and argue that neither of them poses a threat to moral realism.

In Chapter 3, I turn to the argument from explanatory impotence. The proponents of this argument claim that the only facts and properties we are justified in believing are those that are required to explain (i.e. that figure in the best explanation of) our forming the beliefs that we do. According to them, reference to moral facts is not required to explain our moral beliefs. The realist's reply to the argument from explanatory impotence will vary according to the specific criterion of justified belief that is thought to be more appropriate. Some authors claim that the causal criterion of justified belief is too strong, and that the argument from explanatory impotence must turn on the weaker, and more plausible, explanatory criterion. Those who opt for the latter criterion are motivated by the worry that some of our main epistemological achievements (such as mathematical claims or the laws of physics) would fail to meet the former criterion. However, there are also authors who accept the causal criterion and argue that moral facts are causally efficacious. This line of reasoning is generally adopted by those who defend naturalistic versions of moral realism. In Chapter 3, I consider two important objections to the view that moral facts can figure in causal explanations. The first objection holds that moral explanations are simply vacuous. According to the second objection, although moral facts may seem to play causal explanatory roles, it is the nonmoral facts on which moral facts supervene that are actually doing all the causal explanatory work. I argue that moral realism has the resources to meet these objections. Also, I argue that, despite common assumptions to the contrary, the claim that moral facts are causally efficacious can consistently be endorsed by both naturalists and nonnaturalists. In particular, I claim that realists can resist the conclusion that the real causal explanatory work is done by the natural properties on which moral properties supervene by expressing doubts about the truth of the supervenience principle.

The question whether there is reason to believe that there are cases of comparative moral assessment in which the trichotomy of relations "better than", "worse than" and

"equally good" fails will be thoroughly addressed in Chapter 4. In this chapter, I consider at length two of the leading arguments for incommensurability that exist in literature: the argument from multiple rankings and the small-improvement argument. I also address the question of whether value incommensurability should be construed as a determinate failure of the trichotomy of "better than", "worse than" and "equally good" or rather as an indeterminate failure. The argument developed in Chapter 4 is intended to complement an argument from Chapter 2. Whereas in Chapter 2 I argue that comparative indeterminacy poses no threat to moral realism, in Chapter 4 I argue that there are cases of comparative indeterminacy.

In Chapter 5 I address one of the most powerful objections to the view according to which the argument from moral disagreement can be refuted by appealing to the idea of moral indeterminacy. The idea that there are indeterminate moral issues is supposed to account for the possibility of disagreement among agents who suffer from no cognitive shortcoming. This response to the argument from disagreement runs into the following difficulty. Presumably, agents who suffer from no cognitive shortcoming would have to be aware of all morally relevant facts, including those which can render a situation morally indeterminate. And such awareness of the indeterminate character of a moral situation, the objection goes, would undermine the very possibility of disagreement. I argue that moral realism has the resources to answer this objection. In this concluding chapter, I also attempt to tie up two loose ends. First, I look into whether the objectivist who deals with the argument from disagreement would be better served by construing incommensurability according to the model proposed by Joseph Raz or according to the model proposed by John Broome. Second, I attempt to clarify the relationship between value incommensurability and moral dilemmas.

## Chapter 2 The argument from moral disagreement

## 2.1 Moral objectivism

A long-standing argument to the effect that morality fails to be objective has to do with the extent and depth of moral disagreement.<sup>1</sup> The most perspicuous statements of this argument take the form of an argument to the best explanation: the pervasiveness and apparent intractability of moral disagreement are supposedly best explained by denying either that moral statements are the kind of statements which can have a truth-value (according to noncognitivists), or that any moral statements are actually true (according to error theorists).

Before going on to assess the merits of the argument from disagreement, a brief characterization of moral objectivism is in place. Moral objectivism involves embracing both semantic and ontological objectivism.<sup>2</sup> To be a semantic objectivist (cognitivist) about a given area of discourse is to claim that sentences in that area serve to express cognitive attitudes, i.e. they purport to report on how things are in the relevant domain. Given that sentences which lack assertoric force are not fit to express cognitive attitudes, semantic objectivism amounts to the claim that the relevant sentences are assertoric and truth-apt. Semantic objectivism (cognitivism) is a necessary ingredient of an overall objectivist position in ethics. On this view, moral statements express mental states which fall on the cognitive side of the cognitive-noncognitive divide. In other words, moral statements purport to report facts and are apt for truth and falsity.

Moral objectivism involves more than just the view that moral statements are truthapt. It also involves the view that some moral statements are actually true. To claim that sentences in a given area of discourse are truth-apt is to claim that their truth-conditions are well defined. On a cognitivist interpretation of moral discourse, moral statements are apt for truth or falsity in virtue of their representing more or less accurately the facts within the moral domain. However, if there were no moral facts, the truth-conditions of moral statements would never be satisfied, and hence, all moral statements would come out as false. Thus, moral objectivists are bound to hold not only that moral statements purport to report facts, but also that there really are moral facts (ontological objectivism).

Consistently with endorsing semantic and ontological objectivism, one can adopt a relativist view about morality. More concretely, one can hold both that moral statements are truth-apt and that some moral statements are in fact true, but argue that the truthconditions of moral statements are fixed by reference to certain practices or conventions. Yet, no full-fledged objectivist can allow that the truth-conditions of moral statements

<sup>&</sup>lt;sup>1</sup> Proponents of this argument include Ayer (1936: ch. 6), Stevenson (1937; 1948), Hare (1963), Mackie (1977: 36-8), Wong (1984), Williams (1985a: ch. 8), Tolhurst (1987), Harman (1996: ch. 1), Loeb (1998), and Tersman (2006). Replies have been put forward by Wellman (1975), Brink (1984; 1989: 197-210), Hurley (1985), Miller (1985), Boyd (1988), Shafer-Landau (1994; 2003: 215-28), McGinn (1997: ch. 3), Sosa (2001), Wiggins (2005), and Jackson (2008), among others. <sup>2</sup> For this characterization of moral objectivism, see Pettit (2001).

will vary with the practices or conventions that individuals happen to embrace. Moral objectivism is opposed to any such form of relativism.

Given this characterization of what it takes to be a moral objectivist, it follows that there are three ways to be a moral nonobjectivist: (a) embrace a noncognitivist analysis of moral discourse, i.e. hold that moral statements express mental states which fall on the noncognitive side of the cognitive–noncognitive divide, and therefore, are not apt for truth and falsity (*noncognitivism*); (b) embrace the view that moral discourse is undercut by massive error, i.e. hold that, even though moral statements are truth-apt, the facts in light of which such statements could turn out to be true are not to be found (*error theory*); or (c) embrace the view moral statements cannot be assessed for truth and falsity contextindependently (*moral relativism*).

The argument from moral disagreement has been widely relied upon to support each of the above-mentioned nonobjectivist positions. Noncognitivists, error theorists and relativists alike claim that the pervasiveness and apparent intractability of moral disagreement undermine the plausibility of moral objectivism. The main point of contention between them is precisely which of the above-mentioned objectivist tenets has to be rejected.

Henceforth, I will be concerned with establishing that considerations having to do with moral disagreement do not provide a secure basis for rejecting moral objectivism. More specifically, I will argue that moral disagreement can be explained on the basis of a hypothesis which is consistent with the view that moral statements are truth-apt and some moral statements are actually true. Although the arguments presented in this chapter and the subsequent ones have a bearing on whether moral statements should be regarded as being true or false in a straightforward sense (as opposed to being true or false only relative to certain practices or conventions), in order to avoid unnecessarily complicating the discussion further references to the distinction between an objectivist account of the nature of moral truth and a relativist one will be omitted. From now on, by saying that some moral statements are actually true I will mean that they are true in a straightforward sense.

The argument from disagreement holds that the pervasiveness and intractability of moral disagreement are best explained on the assumption that there are no moral facts. The argument goes as follows. If there were moral facts, as the objectivist supposes, then at least one of the parties to a moral disagreement would be subject to a cognitive error. However, given the actual extent of moral disagreement, the most charitable account of the matter is that none of the parties involved needs to be mistaken. Moreover, moral disagreement seems to persist even among competent (i.e. well-informed and rational) inquirers. Had there been any moral facts, we would expect convergence of moral belief among competent inquirers. So it seems that the pervasiveness and apparent intractability of moral disagreement are best explained by denying that there are any moral facts that could be represented by our moral statements.

At this point, it is important to notice two things. First, even if the proponents of the argument from disagreement are right in claiming that moral objectivism has trouble accommodating the distinctive nature of moral disagreement, they must also establish that their own explanation of disagreement is superior to the one implying the existence of moral facts. Second, granting that the best explanation of moral disagreement would involve denying that moral facts exist, further argument would be needed to establish that

disagreement is to be explained along the lines of noncognitivism, error theory or relativism. Nevertheless, in what follows I will mainly be interested in establishing that moral objectivism can accommodate the distinctive nature of moral disagreement, and not in coming up with a comparative assessment of how objectivism and its competitors explain disagreement.

The preliminary characterization of moral objectivism outlined above leaves an important question unanswered. Objectivists share the view that some moral statements are actually true. However, they disagree about what it is that makes moral statements true, i.e. they disagree about the nature of moral facts. Now, the question that arises is whether an appropriate conception of moral objectivity would require claiming that the existence and nature of moral facts is mind-independent. Answering this question is especially important for my present purposes, since it will help establish what sort of considerations may legitimately be invoked by moral objectivists in responding to the argument from disagreement. Let me detail.

In characterizing moral cognitivism it is customary to distinguish between realist and constructivist theories. According to moral constructivists, the truth-conditions of moral statements are fixed by reference to the deliverances of some idealized agent or group of agents.<sup>3</sup> On this view, a moral statement is true if and only if it accurately reflects the deliverances of that agent or group of agents. In other words, there is no moral truth prior to or independent of the deliverances of such agents. By contrast, moral realists claim that whether a moral statement is true is to be determined independently of the deliverances of any actual or hypothetical agent.

Having distinguished between moral realism and moral constructivism, the question is whether constructivist theories can properly be counted as objectivist. Before answering this question, however, it will be useful to consider the question whether all versions of realism can properly be counted as objectivist.

In contrast with the constructivist view according to which moral facts are constituted by, or constructed from, the responses of some ideal agent or group of agents, moral realism holds that moral facts are *robust*, i.e. it holds that the metaphysical status of moral facts is not relevantly different from that of certain ordinary nonmoral facts.<sup>4</sup> Moral realism comes in different versions, not all of which claim that the instantiation of moral properties is entirely independent of characteristic human responses. According to the so-called "sensibility theories" or "response-dependent theories", moral properties enjoy a

<sup>&</sup>lt;sup>3</sup> See, for instance, Firth (1952), Rawls (1980), Scanlon (1982), and Milo (1996). Compare relativist versions of constructivism, according to which the truth-conditions of moral statements are fixed by reference to certain practices or conventions. For the latter, see Harman (1975; 1996) and Wong (1984).

<sup>&</sup>lt;sup>4</sup> For this characterization of moral realism, see Väyrynen (2006). According to Väyrynen, attributing to moral realists the view that moral facts and properties are metaphysically robust "affords one (albeit not the only) way of capturing many realists' conviction that ethics concerns objective matters of fact whose existence and nature are independent of anyone's sentiments, opinions, evidence, or theories about what is right or wrong" (2006: 380). It should be emphasized, however, that the definition of moral realism in terms of mind-independence is not extensionally equivalent to Väyrynen's definition. Since the idea of mind-independence admits of different interpretations, establishing which is the relevant interpretation is essential. Yet, if a stronger interpretation of mind-independence is taken as relevant, then the definition of moral realism in terms of mind-independence would classify as irrealist certain views that count as properly realist on Väyrynen's definition. For more on this, see note 5 bellow.

status similar to the one enjoyed by secondary qualities.<sup>5</sup> Secondary qualities of objects are dispositions to elicit certain characteristic human responses. As dispositions to cause certain human responses, they are persisting properties of objects (i.e. these dispositions persist even when they are not triggered). So, if moral properties are similar to secondary qualities, then there is a legitimate sense in which they are real properties of objects. Another way of construing this point would be to say that, although moral properties can achieve salience only in virtue of causing certain responses in us, they do not depend for their existence on any particular human response.

If morality has something essentially to do with agents and their sensibilities, then it falls short of a kind of objectivity. It is important to notice, however, that this is not the kind of objectivity that is at stake when discussing the argument from disagreement. The advocates of this argument hold that moral disagreement cannot be explained consistently with the view that there are moral facts. If an explanation of disagreement consistent with a broadly anthropocentric view of the nature of moral facts were available, the argument from disagreement would be refuted. In fact, some of the advocates of the argument from disagreement claim that:

The way in which moral practices vary with the forms of life of a society is not at all similar to the way, if any, in which perceptions of secondary qualities can vary with those forms of life. [...] Apart from rare borderline cases, there is nothing in the secondary quality ascription parallel to the 'essentially contested' character of many moral verdicts. (Blackburn, 1985*a*: 15)

It follows, then, that the sense of objectivity that is at stake when discussing the argument from disagreement is not the sense in which objectivity is the opposite of subjectivity, but the sense in which it is the opposite of nonobjectivity. Therefore, in responding to the argument from disagreement, moral objectivists may legitimately invoke considerations that derive from a broadly anthropocentric view of the nature of morality. As Philip Petit points out:

The construal of objectivity as the opposite of subjectivity may suggest that someone who espouses ethical objectivism in any full sense of the term must be opposed to a broadly subjective or anthropocentric view of the values and disvalues

<sup>&</sup>lt;sup>5</sup> The main proponents of this view are McDowell (1985) and Wiggins (1987c). On the definition of moral realism proposed here, sensibility theories count as realist, for the metaphysical status of moral properties is not relevantly different from the one enjoyed by other ordinary nonmoral properties (i.e. colors). However, other definitions of moral realism classify sensibility theories as irrealist. See, for instance, Dancy (1986), for an argument to the effect that any view which claims that moral properties are constituted by the availability of a characteristic human response is incompatible with moral realism. (Some might argue that sensibility theories count as irrealist even on Väyrynen's definition. Admittedly, the analogy between moral properties and colors is far from unproblematic. Moreover, according to Väyrynen (2006), one way to elucidate the idea that certain types of properties are metaphysically robust is to say that such properties figure in causal explanations. However, the argument would go, moral properties (or, at least, moral properties as sensibility theorists conceive of them, i.e. as nonnatural properties) seem to be causally inert. I believe that this line of reasoning is misguided. In Chapter 3, I will argue that moral properties are causally efficacious, and that, despite common assumptions to the contrary, this claim can consistently be endorsed by both naturalists and nonnaturalists. The fact that sensibility theories have the resources to respond to the argument from explanatory impotence is, in my view, reason enough to count such theories as properly realist.)

countenanced. But this consideration ought to have no hold on us. [...] Ontological objectivism will be vindicated by the reality of any ethical values, even values that that are decidedly immanent to human life [...]. (Pettit, 2001: 244-5)

Now, let me turn to the question whether moral constructivism can properly be counted as objectivist. Recall that relativist versions of constructivism allow for there being a plurality of sets of moral facts identified by reference to different practices or conventions. By contrast, constructivists that adopt ideal agent theories claim that there is a *single* set of moral facts, and therefore, I believe that they should not be excluded out of hand from the objectivist camp.<sup>6</sup> In what follows, I will mainly be concerned with realist versions of cognitivism. However, in Chapter 5 I will consider a response to the argument from moral disagreement which is available only to constructivists. If it turns out that this is the only adequate response to the argument from disagreement, this might motivate a shift from moral realism to (nonrelativist) moral constructivism.

#### 2.2 Noncognitivism and error theory

Although both noncognitivists and error theorists argue that the distinctive nature of moral disagreement is best explained by denying that morality is objective, the two lines of argument are incompatible. One cannot consistently hold both that moral judgments are not apt for truth and falsity and that moral judgments turn out to be false. Nonetheless, considerations deployed to defend each of the aforementioned views might be used to support the other. As Geoffrey Sayre-McCord points out:

[S]omeone defending an error theory might point to the ways in which moral claims are used to express or serve peoples' emotions, attitudes, and interests, to explain why people keep arguing as they do despite there being no moral facts. And someone defending noncognitivism might point to the practical utility of talking as if there were moral facts to explain why moral claims seem to purport to report facts. (Sayre-McCord, 2005)

However, a closer look at the reasons advanced by the proponents of each of these views for endorsing their favored account of disagreement will also reveal the difficulties that beset error theory, on the one hand, and noncognitivism, on the other. Consider, for instance, John Mackie's view according to which moral discourse is undercut by massive error. Mackie's case against moral objectivism boils down to two main arguments, i.e. the argument from queerness and the argument from disagreement. Roughly put, the idea behind the argument from queerness is that moral objectivists are committed to dubious metaphysical and epistemological claims. Objective moral facts, if there were any, would

<sup>&</sup>lt;sup>6</sup> The question whether nonrelativist versions of constructivism can properly be counted as objectivist has received different answers in the literature. For instance, according to Sayre-McCord (1988*a*), contractarian versions of constructivism count as intersubjectivist, whereas ideal observer theories count as subjectivist. This may seem odd, especially since Sayre-McCord is willing to count as properly realist all versions of cognitivism with the exception of error theory. Compare Shafer-Landau (2003).

be facts of a very strange sort, "utterly different from anything else in the universe" (Mackie, 1977: 38). In addition to their being metaphysically extraordinary, objective moral facts would also require an extraordinary epistemology. If we had epistemic access to such facts, it would have to be through a special faculty of moral intuition. However, we are left in the dark as to how the exercise of this faculty differs from our ordinary ways of knowing everything else. I will consider in detail a version of the argument from queerness in the following chapter. At this point, what is important to notice is that, even if each of the aforementioned two arguments provided a secure basis for rejecting moral objectivism, neither of them would favor error theory over noncognitivism. One can agree with Mackie that moral objectivism is a metaphysically queer doctrine and that it fails to explain the persistence of moral disagreement, and still argue for a noncognitivist account of moral discourse.

Mackie's reasons for not siding with noncognitivism have to do with the fact that, in his view, both emotivist and prescriptivist theories fail to do justice to the basic meaning of ethical terms. He points out that:

The ordinary user of moral language means to say something about whatever it is that he characterizes morally, for example a possible action, as it is in itself, or would be if it were realized, and not about, or even simply expressive of, his, or anyone else's, attitude or relation to it. (Mackie, 1977: 33)

Mackie goes on to stress that, when someone is in a state of moral perplexity, wondering, for instance, whether to engage in research connected to bacteriological warfare, the question is not "whether he really wants to do this work, whether it will satisfy or dissatisfy him, whether he will in the long run have a pro-attitude towards it, or even whether this is an action of a sort that he can happily and sincerely recommend in all relevantly similar cases" (ibid.). Someone confronted with such a choice wants to know whether engaging in work connected bacteriological warfare would be wrong in itself. According to Mackie, any analysis of the meaning of moral statements which leaves out the claim to objective validity or truth is incomplete. It is precisely at this point that Mackie parts company with noncognitivists. In fact, if second order ethics were confined to analyzing the meaning of moral statements, Mackie would have no quarrel with objectivists. Given that, according to him, the claim to objectivity is ingrained in our thought and built into the way in which moral language is used, any argument against moral objectivism could only impact on the truth-value of moral statements, and not on their truth-aptness.

Now, let me turn to the reasons advanced by noncognitivists for rejecting error theory. Interestingly enough, a common worry about Mackie's view is shared by several of his critics. For instance, Blackburn (1985*a*: 1–3) notices that Mackie does not draw the consequences one might expect from his error theory.<sup>7</sup> Although in the first chapter of his *Ethics: Inventing Right and Wrong* Mackie claims that ordinary moral discourse is infected by error, in the second part of the book he deals with a large number of substantive issues, using the same, supposedly erroneous, moral concepts. Mackie (1977: 49, 105–6) does consider the question whether adopting an error theory rules out all normative ethics, and is aware that some will be inclined to answer this question in the

CEU eTD Collection

<sup>&</sup>lt;sup>7</sup> See also Williams (1985*b*).

affirmative. Yet, he holds that even though the claim to objectivity which is part of the very meaning of moral concepts is unwarranted, there is enough room left for first-order moralizing. In Humean spirit, he points to the socially useful function of morality: "Morality is not to be discovered but to be made: we have to decide what moral views to adopt" (1977: 106). However, Blackburn argues that:

[F]rom the standpoint of an error theory, it is quite extraordinary that we should have to do any such a thing. Why should we have to choose to fall into error? Surely it would be better if we avoided *moral* (erroneous) views altogether and contented ourselves with some lesser, purged commitments that can be held without making metaphysical mistakes. (Blackburn, 1985*a*: 2, emphasis in original)

That Mackie goes on to advocate various straightforward moral views is, according to Blackburn, "enough of a puzzle to cast doubt back on to the original diagnosis of error" (ibid). Blackburn points out that if a given vocabulary embodies an error, then it would be better if it were replaced by a different vocabulary which avoids the charge. Thus, if the *mere use* of our moral concepts embodies an error, then it would be better to substitute for them a new set of concepts which serve our legitimate interests but avoid the old error. That Mackie provides no clue about how such an error-free vocabulary would look like suggests that no error is incorporated in the *mere use* of moral concepts.

Blackburn considers a possible reply to this objection. A proponent of Mackie's position could argue that the meaning of the terms used within a certain area of discourse largely depends on the theory one holds about the nature of the discourse in question. Such holistic considerations about how our theories infect our meanings can be appealed to in order to claim that, although the ordinary meaning of moral terms embodies an error, Mackie rids himself of the error by endorsing a different theory about the nature of moral discourse. On this view, the 'full meaning' of moral terms is determined by both the practice of moralizing and whatever theory the subjects hold about the nature of this practice. The holist will then have it that the full meaning of the terms employed by Mackie is altered by his adopting a Humean picture of the nature of morality. Hence, there would be no tension between advocating an error theory, on the one hand, and continuing to moralize, on the other.

However, Blackburn argues that it is in principle possible to scrutinize the practice of some subjects without being able to tell which theory the subjects hold about the nature of their practice. The holist will have it that those who endorse a correct (Humean) theory about the nature of moral practice give a different full meaning to the terms they employ. Yet, Blackburn rightly points out that, since the practice of those who hold a Humean theory is no different from the practice of those who fall prey to the objectivist error, it cannot follow that the *practice* of those who hold the wrong theory – as opposed to the practice of those who hold a correct theory – embodies an error. The conclusion that suggests itself is rather that practice, as such, is error-free. Hence, the holist has to give up Mackie's claim that the practice of ordinary moralists embodies an error.

Although I agree with Blackburn that there is something odd about advocating an error theory and yet continuing to moralize, I find the last step in his argument objectionable. According to the 'holistic' defense of Mackie's position, whether error is embodied in the practice of some subjects will directly depend on whether the theory they

hold about the nature of their practice is erroneous – hence, there is no point in dealing with these two issues separately. Against the holist, Blackburn argues that since practice is in fact the same regardless of what theory one holds about its nature, we are led to the conclusion that, on the one hand, there is an error-free practice, and on the other hand, a variety of (possibly erroneous) theories about its nature (1985*a*: 3). This conclusion seems to be tailored specifically for Blackburn's purposes. For Blackburn wants to claim that despite Mackie's being right about the metaphysical issue, ordinary moral discourse embodies no error. If Blackburn's noncognitivist analysis of moral discourse is correct (i.e. if the claim to objectivity is not part of the very *meaning* of moral terms), then the mere use of our ordinary moral terms cannot *possibly* involve an error. However, there are reasons to doubt that Blackburn's argument against the holist does establish that practice, as such, is error-free. Let me detail.

The holist supposes that we can distinguish between a practice that embodies error and an error-free practice on the basis of the theory the subjects hold about the nature of their practice. If Blackburn is right in claiming that the practice of those who hold an erroneous theory is, in fact, no different from the practice of those who hold a correct theory, then the holist has to give up the claim that error at the level of theory and error at the level of practice are intimately related. Hence, the holist cannot argue that the tension between Mackie's advocating an error theory, on the one hand, and his continuing to moralize, on the other, is only apparent. This is because the holist has to give up the claim that the practice of those who hold an erroneous (objectivist) theory -as opposed to the practice of those who hold a correct (Humean) theory – embodies an error. At this point, however, it is not at all clear whether the holist will have to admit that (a) the practice of those who hold an erroneous theory is error-free (and hence, Mackie's error theory should be rejected), or that (b) from the standpoint of an error theory, all practice embodies error (and hence, there is something awkward about the fact that Mackie continues to moralize). According to Blackburn, the holist will have to admit (a). Now, we should keep in mind that Blackburn's argument against the holist is premised upon the assumption that a correct theory about the nature of moral practice is one which points to its strictly instrumental value. Insofar as Mackie is right about the metaphysical issue, the practice of those who hold an erroneous (objectivist) theory about the nature of moral practice can only be error-free if the claim to objectivity is not part of the very meaning of the terms they use. Admitting (a) would thus be tantamount to dropping Mackie's claim that ordinary moral terms include, as part of their very meaning, a claim to objectivity. Hence, reaching the conclusion (a) would work (although indirectly) to the credit of noncognitivists. By contrast, reaching the conclusion (b) would be congenial to cognitivists. For cognitivists will readily admit that, once we assume that Mackie is right about the metaphysical issue, we cannot escape the conclusion that ordinary moral practice embodies error.

But is it really the case that the holist will have to accept (a)? If practice is in fact the same regardless of what theory one holds about its nature, it cannot be the case that the practice of those who hold an erroneous theory – *as opposed* to the practice of those who hold a correct theory – embodies error. So far, no reason has been provided for accepting (a) rather than (b). Then, why does Blackburn conclude that all practice is error-free? Blackburn's point seems to be that, since observing the practice of some subjects leaves us unable to tell which theory they hold, we can *never* rule out the

possibility that the subjects hold the correct theory. However, this argument works just as well in the opposite direction. For all we know, the subjects may as well hold the wrong theory. If this is also a possibility that can *never* be ruled out, then why not conclude that all practice embodies error?<sup>8</sup>

The upshot of all these considerations is that endorsing Blackburn's argument against the holist does not commit us to rejecting Mackie's claim that ordinary moral terms include, as part of their very meaning, a claim to objectivity. In other words, holding that there is a tension between advocating an error theory and yet continuing to moralize does not commit us to Blackburn's noncognitivism.

Before going further, let me briefly consider another possible objection to Blackburn's argument. This objection would go as follows. According to Blackburn, we could closely observe the practice of some subjects and "know as much as there is to know about their ways of thinking, commending, approving, deliberating, worrying, and so on, and yet be unable to tell from all that which theory they hold" (1985*a*: 3). This is a controversial claim. Given that realists often defend their view on phenomenological grounds, they will have difficulties in admitting that the way we conduct ourselves in moral inquiry can be reconciled both with a cognitivist account of moral discourse and with a noncognitivist account. Blackburn holds that practice is in fact the same regardless of what theory one holds about its nature – the objection would go – precisely because he is confident that his own version of noncognitivism can accommodate the realist-seeming features of moral discourse.

This objection relies on a misinterpretation of the realist's position. Those who appeal to considerations about the way we conduct ourselves in moral inquiry in order to defend moral realism hold that moral phenomenology is most easily explained on realist assumptions. Nonetheless, when realists emphasize that the fact that we engage in moral debates and that we regard ourselves and others as capable of making moral mistakes is best understood on realist assumptions, they do not claim that belief in moral realism is a *common belief.*<sup>9</sup> The interest of realists lies with the philosophical implications of moral thought and inquiry, not with whether people *actually* uphold certain metaethical views. Hence, realists have no reason to object to Blackburn's claim that moral practice is in fact the same regardless of what theory the subjects hold about the nature of their practice.

Now, let me rehearse some of the points made so far. On the one hand, Mackie argues that ordinary moral terms include, as part of their very meaning, a claim to objectivity. Although he shares with noncognitivists a conception of the world as devoid of any moral facts, he contends that any noncognitivist analysis of moral discourse is faulty because it fails to do justice the basic meaning of moral terms. On the other hand, noncognitivists point out that there is something awkward about holding both that ordinary moral discourse embodies an error and that no radical revision is called for by admitting that fact. Noncognitivists find it odd to "insist on construing the discourse in

<sup>&</sup>lt;sup>8</sup> A proponent of Blackburn's argument can reply that, even if the possibility that the subjects hold an erroneous theory can never be ruled out, it cannot follow that *all* practice embodies error, since we have assumed from the outset that there are subjects who hold a correct theory. It should be noticed, however, that Blackburn's own argument does not rely on denying that there are subjects who hold the wrong theory. It is our being unable to *identify* the subjects who hold the wrong theory that leads Blackburn to conclude that all practice is error-free. Yet, since we are equally unable to identify the subjects who hold the correct theory, why not hold that all practice embodies error?

<sup>&</sup>lt;sup>9</sup> For this point, see Brink (1989: 25).

terms of a notion of truth which has us in massive error", when the alternative of construing moral discourse in noncognitivist terms would avoid the difficulties into which Mackie's error theory collapses.<sup>10</sup>

My aim in this section was to offer a brief picture of the shortcomings of noncognitivism, on the one hand, and of error theory, on the other. However, as it is often pointed out, considerations against one metaethical view or another are seldom decisive, being primarily meant to establish where the burden of proof lies.<sup>11</sup> So even if noncognitivism does not do justice to moral phenomenology<sup>12</sup>, noncognitivists can argue that there are other grounds for concluding that the most defensible metaethical view is the one that they advocate.<sup>13</sup> Moreover, a full assessment of the merits of noncognitivism would also require assessing the merits of its more sophisticated versions that seek to vindicate our ordinary understanding of morality.<sup>14</sup> Likewise, even though many find unpalatable the error theorist's claim that there is something essentially misguided about our moral practice, this constitutes no conclusive proof that error theory must be rejected.

More importantly, however, leaving aside the state of current debates between noncognitivism and error theory, I want to make the following point. When we look at the main arguments that noncognitivists and error theorists use against each other, we notice that they usually come down to mutual accusations of going against common sense. Underlying this strategy seems to be an attempt to preserve as many intuitions as possible from those typically adduced in favor of moral realism. Now, since even the opponents of moral realism appeal to considerations that typically motivate a realist stance, this reinforces the idea, already adumbrated in the Introduction, that realism is the default position in metaethics. In other words, we should only be led to adopt other views if the metaphysical and epistemological commitments underlying realism cannot be successfully defended. This by itself does not necessarily strengthen the case for moral realism, but it does however provide an incentive for more careful analysis into the resources available to realists for responding to their opponents' attacks. It is this observation that motivates the direction of my present research, which is geared towards assessing the argument from moral disagreement and the potential of moral realism to keep this argument at bay.

<sup>&</sup>lt;sup>10</sup> Wright (1992: 87). Wright explicitly claims that the most important question concerning Mackie's error theory does not have to do with its coherence, but with its *motivation*. Compare Blackburn (1985*b*) and Williams (1985*b*).

<sup>&</sup>lt;sup>11</sup> See Shafer-Landau (2003: 3).

<sup>&</sup>lt;sup>12</sup> The term is used here in a wide sense, to include both facts about the way we conduct ourselves in moral thought and moral inquiry, and semantic facts about the form and content of our ordinary moral language.

<sup>&</sup>lt;sup>13</sup> Admittedly, noncognitivists do not base their view on considerations that have to do with moral phenomenology. The main motivation for noncognitivism has traditionally lied with the belief that cognitive attitudes are motivationally inert.

<sup>&</sup>lt;sup>14</sup> See, for instance, Blackburn (1993). Blackburn's "quasi-realist" project places a special emphasis on explaining why moral statements can be regarded as truth-apt even though strictly speaking they are neither true nor false. See also Gibbard (1990).

### 2.3 Objectivity and convergence

The most common version of the argument from moral disagreement has as its premise the existing diversity of moral belief.<sup>15</sup> The proponents of this argument claim that the pervasiveness and apparent intractability of moral disagreement are best explained by dropping the assumption that there are moral facts. If there were moral facts, at least one of the parties to a moral disagreement would be subject to a cognitive error. Yet, given that moral disagreement is so pervasive, the most charitable account of the matter is that none of the parties involved needs to be mistaken.<sup>16</sup> Moreover, disagreement seems to persist even among competent (i.e. well-informed and rational) inquirers. According to the proponents of the argument from disagreement, if there were any moral facts, we would expect convergence of moral belief among competent inquirers at least over time. The fact that disagreement often concerns moral issues that have been extensively debated by competent inquirers is taken to entail that there is no fact of matter concerning the issues over which they disagree.

A commonplace among those who oppose the argument from disagreement is to point out that from the fact that there is persistent disagreement in the natural or social sciences we do not infer that these areas of inquiry do not deal with matters of objective fact.<sup>17</sup> Scientists may continue to disagree about the origin of our species or the existence of Atlantis, but we do not take this as evidence that there is no fact of the matter regarding these issues. So the proponents of the argument from disagreement must establish that there is something peculiar about moral disagreement that poses a serious threat to moral objectivity.

One way to bring out the contrast between moral and scientific disagreement is to point out that the former is more pervasive than the latter. According to the advocates of the argument from disagreement, the pervasiveness of moral disagreement is best explained on the assumption that there are no moral facts. For instance, Mackie writes that:

[R]adical differences between first order moral judgments make it difficult to treat those judgments as apprehensions of objective truths. [...] In short, the 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. (Mackie, 1977: 36-7)

Mackie's argument is mainly directed against moral intuitionism, which holds that moral truths can be apprehended directly through a special faculty of moral intuition.<sup>18</sup> According to him, moral intuitionism is rendered implausible by the pervasiveness of moral disagreement. In order to account for moral disagreement, the intuitionist has to invoke the idea of a distorted perception of the moral truth. Yet, how can it be that our special intuitive faculty goes awry in so many cases? Surely, Mackie argues, a charitable

<sup>&</sup>lt;sup>15</sup> A different version of the argument holds that the mere fact that disagreement among fully competent inquirers is possible gives us reason to reject moral realism. See Tolhurst (1987) and Wright (1992).

 <sup>&</sup>lt;sup>16</sup> See, for instance, Mackie (1977: 36–7).
<sup>17</sup> See Brink (1989: 198–9), Pettit (2001: 255), and Shafer-Landau (2003: 7, 220).

<sup>&</sup>lt;sup>18</sup> For a discussion of Mackie's argument, see Loeb (1998: 282–5).

and more plausible account of widespread discrepancies in moral belief is that they reflect different ways of life.

Although moral intuitionism seems to have lost much of its appeal among contemporary philosophers, Mackie's treatment of the argument from disagreement provides an important insight. Why does Mackie claim that the best explanation of the pervasiveness of moral disagreement is that moral questions do not admit of objectively true answers? The underlying thought is that, if there are moral truths, the failure of these truths to gain recognition among a significant number of persons is in need of explanation. In the absence of a convincing explanation of why such truths would impinge on some but not on others, we are led to the conclusion that there are no objective moral truths. According to Mackie, moral intuitionism falls short of providing such an explanation. For there is something odd about holding both that we have a special intuitive faculty for apprehending moral truths and that this faculty goes wrong so often.

Now, those who oppose the argument from disagreement can point out that the tenability of moral realism does not depend on the assumption that we have direct access to moral truths through a special intuitive faculty. Yet, the argument from disagreement remains in force with regard to any view which holds that there are objective moral truths and that we have epistemic access to these truths. Needless to say, the mere occurrence of disagreement does nothing to undermine the claim to objectivity. Disagreements occur also in areas of inquiry where we are confident about objectivity, and we do not hesitate to attribute them to the parties' cognitive shortcomings. It is the pervasiveness of moral disagreement that prompts a worry about moral objectivity. In order to hold on to the idea that there are moral truths, moral realists have to claim that at least one party to a moral disagreement is either ill informed or ill endowed to judge the issue under consideration, or subject to some sort of bias. However, given that moral questions elicit extensive disagreement, moral realists have to claim that a significant number of persons are so constituted or so positioned as to fail to answer such questions correctly. The advocates of the argument from disagreement find this last claim implausible. They commonly support their view by emphasizing that it is precisely in areas where we witness a great deal of agreement that we feel reassured about objectivity (e.g. natural sciences and mathematics), whereas lack of agreement in an area usually makes us question objectivity (e.g. esthetics). This places the burden of proof on the moral realist, i.e. it is the moral realist who has to show that the failure to reach agreement on moral issues is better explained by the hypothesis that at least one party to a moral disagreement is subject to a cognitive shortcoming than by the hypothesis that there are no objective moral truths.

According to the proponents of the argument from disagreement, the threat to moral objectivity does not only come from the pervasiveness of moral disagreement, but also from the fact that many moral disputes seem to resist rational resolution. The proponents of this argument typically hold that realism about a domain entails that all the disputes associated with the domain in question are in principle resolvable. The fact that many moral disagreements are resistant to rational resolution, the argument goes, warrants a rejection of moral realism. Fully fleshed out, this argument requires the supposition that moral disagreements are irresoluble not only in practice, but also in principle. More specifically, it requires the supposition that moral realists cannot explain the lack of resolution in practice consistently with resolution being in principle available. Let me detail.

The failure to reach a consensus would pose no threat to moral objectivity if it were clear that at least one of the parties to a moral disagreement were ignorant of the relevant nonmoral facts or irrational. This is because the claim that disagreement stems from the parties' cognitive shortcomings amounts to the claim that, even though disagreement is not resolved in practice, it is in principle resolvable. However, the hard fact is that wellinformed and rational persons may continue to disagree on moral issues even after prolonged reasoning with one another. According to the advocates of the argument from disagreement, the fact that even prolonged reasoning fails to lead well-informed and rational persons to moral consensus shows that disagreement about the issue in question is in principle irresoluble. How plausible is this claim?

Those who oppose the argument from disagreement commonly emphasize that we are reluctant to take persistent disagreement among competent scientists as a reason for inferring that there is no right answer to the question being addressed. Then why should we suppose that the persistence of moral disagreement rules out the claim to objective truth? The opponents of the argument from disagreement point out that "there is always a logical gap between the result of actual reasoning among a limited number of imperfectly rational men through a limited period of time and the ideal agreement projected by the claim to truth" (Wellman, 1975: 212). On the other hand, the proponents of the argument from disagreement maintain that, whereas scientific disagreements can be attributed to the parties' cognitive shortcomings (e.g. to lack of adequate evidence), in what moral disagreements are concerned, such explanations seem, all too often, misplaced.<sup>19</sup> Insofar as the proponents of the argument from disagreement are right in claiming that moral disagreement need not involve anything worth regarding as a cognitive shortcoming, we are led to the conclusion that moral disputes are in principle irresoluble. Thus, in order to refute the argument from disagreement, the realist must prove either that (a) every moral disagreement involves a cognitive shortcoming, or that (b) the fact that some moral disputes are in principle irresoluble poses no threat to moral realism. Addressing the question of how well can the realist accomplish this task is the main aim of the next section.

In what follows, I will focus on a different question. Just what is the connection between there being moral facts and convergence of moral belief? Admittedly, the answer to this question has to do with epistemological considerations.<sup>20</sup> If there were any moral facts, then those who are not subject to any cognitive shortcoming would obtain knowledge of these facts. Thus, if realists are right in claiming that there are moral facts, we have reason to expect convergence of moral belief among competent inquirers at least over time.

The failure to achieve convergence of moral belief may signal the fact that some (or all) of the parties involved are subject to a cognitive shortcoming. Yet, it may just as well indicate that there is no moral reality to which the parties' beliefs can be answerable. The more often the realist has to invoke the cognitive shortcomings of the parties in order to explain disagreement, and the more mysterious the nature of these shortcomings, the less

 <sup>&</sup>lt;sup>19</sup> See, for instance, Mackie (1977: 36).
<sup>20</sup> For more on this, see Tersman (2006: 52ff).

plausible is the view that there are moral facts. This is a point that realists are willing to admit. Consider, for instance, the following passage from David Wiggins:

Suppose I am convinced that something is so. Then it is disturbing to me if nobody else can be brought to agree with me. Why? Well, if something is so *either* it must be capable of impinging on others in the way it impinged on me or I shall have in principle to account for its inaccessibility to all others. And if I could have accounted for that, then I should never have been disturbed in the first place by disagreement. If however there were no prospect at all that arguments founded in what made me think it true should have non-random efficacy in securing agreement about whether p, I should be without protection from the idea that [...] there was just nothing at issue. (Wiggins, 1987b: 149, emphasis in original)

The idea underlying this passage is that, insofar as the realist holds that we have epistemic access to moral facts, he must provide an explanation of why these facts are accessible to some but not to others. Note that to claim that one *knows* that p is to claim not only that one believes that p and that p is true, but also that one is justified in believing that p. And saying that one is justified in believing that p means that one has adequate evidence that p. Now, if a person can come to know that p, i.e. if that person has good evidence that p, we would expect that other persons would also posses that evidence and believe that p.<sup>21</sup> In other words, we would expect (a measure of) convergence on the belief that p. If, however, there were no prospect to achieve convergence of moral belief among "those not disabled from grasping the truths of this subject matter", we would have to conclude that there are no moral facts.<sup>22</sup>

By contrast, the existence of convergence of moral belief would provide support for the view that there are moral facts. As John Skorupski emphasizes:

If individuals with 'different characters' converge on the same judgment, the judgment is so much the less likely to be explained by the quirks, obsessions, or partialities of an individual temperament. (Skorupski, 1999: 11)

Skorupski goes on to stress that:

[S]table and spontaneous agreement indicates the objective validity of our judgment-making dispositions, by signaling that we are tracking the same, judgment-independent, states of affairs. Convergence of our judgments betokens that they correspond to, or that they co-respond to, a fact. (Skorupski, 1999: 12–3)

The realist can resist the claim that the existence of moral facts gives us reason to expect convergence in ethics by insisting that (at least some) moral facts are undetectable. Admittedly, there is no reason to suppose that any truth is accessible to human inquiry. For instance, no human can know exactly how many stars are in the universe, although there is a right answer to this question.<sup>23</sup> Yet, whereas the claim that truth can outstrip our cognitive capacities may be plausible in some areas of inquiry, it seems unlikely in ethics. Given that ethics is mainly concerned with how people ought to treat each another, it is

<sup>&</sup>lt;sup>21</sup> For emphasizing this point, see Gowans (2000: 16).

<sup>&</sup>lt;sup>22</sup> The quote comes from Wiggins (1990–91: 66).

<sup>&</sup>lt;sup>23</sup> See Sosa (2001: 217).

rather odd to hold that moral truths cannot be known.<sup>24</sup> It seems then that claiming that moral truths are undetectable would only be an *ad hoc* move that would allow the realist to eschew a troublesome argument.

Another strategy available to the moral realist in responding to the argument from disagreement is to point out that, even though we disagree on many moral issues, there is also a great deal of moral consensus. As Pettit rightly emphasizes, although we disagree about the justifiability of capital punishment or about the justifiability of treating women in a certain manner, "such disagreements are quite consistent with agreement on deeper matters, like the justice of treating similar cases similarly or of allowing no one arbitrary, unchallengeable discretion over the destiny of another" (2001: 255). Furthermore, realists stress that there has been considerable moral progress over time.<sup>25</sup> They draw attention to the fact that many moral issues that used to divide the public (e.g. slavery, female voting or interracial marriage) are no longer controversial. At this point, it is important to recall that the argument from disagreement relies on a disanalogy between moral and scientific inquiry. According to this argument, it is the contrast between the degree of convergence that we witness in moral inquiry, on the one hand, and scientific inquiry, on the other, that undermines the plausibility of moral realism. As already pointed out, moral realists typically respond that there is a considerable amount of agreement on moral issues. Also, they suggest that the impression that disagreement on moral issues is more widespread than disagreement in other areas may have to do with the fact that, given their practical implications, moral issues are at the forefront of public consciousness.<sup>26</sup>

Now, let us grant for the moment that the realist is right in claiming that there is a significant amount of convergence of moral belief and ask whether this would be enough to vindicate the view that there are moral facts. Bernard Williams answers this question in the negative. He writes that:

The basic idea behind the distinction between the scientific and the ethical, expressed in terms of convergence, is very simple. In a scientific inquiry there should ideally be convergence on an answer, where the best explanation of the convergence involves the idea that the answer represents how things are; in the area of the ethical, at least at a high level of generality, there is no such coherent hope. The distinction does not turn on any difference in whether convergence will actually occur, and it is important that this is not what the argument is about. It might well turn out that there will be convergence in ethical outlook, at least among human beings. The point of the contrast is that, even if this happens, it will not be correct to think it has come about because convergence has been guided by how

<sup>&</sup>lt;sup>24</sup> For instance, Nagel writes that: "The connection between objectivity and truth is closer in ethics than it is in science. I do not believe that the truth about how we should live could extend radically beyond any capacity we might have to discover it." (1986: 139)

<sup>&</sup>lt;sup>25</sup> See Brink (1989: 208) and Smith (1994: 188). For an interesting argument about why the fact that moral progress has been less impressive than scientific progress poses no threat to moral realism, see McGinn (1997: 46–7).

<sup>&</sup>lt;sup>1</sup>(1997: 46–7). <sup>26</sup> See Wellman (1975: 212). See also Nagel, for stressing that there is considerable disagreement about social and scientific facts, "especially where strong interests are involved which will be affected by different answers to a disputed question" (1986: 148). Nagel claims that it is the presence of such interests that explains why moral disagreements typically receive more attention. He writes that "for comparably motivated disagreements about matters of fact, one has to go to the heliocentric theory, the theory of evolution, the Dreyfus case, the Hiss case, and the genetic contribution to racial differences in I.Q." (ibid.).

things actually are, whereas convergence in the sciences might be explained in that way if it does happen. (Williams, 1985*a*: 136).

In short, the idea is that convergence of moral belief might be explained by a hypothesis that does not involve the assumption that there are moral facts (for instance, it might be explained by reference to our shared upbringing). Thus, the existence of convergence on moral issues would not automatically lead to the conclusion that there are moral facts. Williams's suggestion is that convergence of scientific belief, *as opposed to* convergence of moral belief, is the result of a reliable epistemic mechanism. His argument calls for a story about how we can gain epistemic access to moral facts. So it seems that, in order to conclusively establish that there are moral facts, one would need to show both that there is a considerable amount of convergence of moral belief, and that this convergence of moral belief can suitably be regarded as the result of a reliable epistemic mechanism.

At this juncture, those who attempt to answer the argument from disagreement can draw attention to the following point. The idea underlying (the most common version of) the argument from moral disagreement is that agreement in an area is a reliable indicator of objectivity. More specifically, the idea is that insofar as competent (i.e. rational and well-informed) inquirers converge in their beliefs pertaining to facts in a given area, we have reason to accept the objectivity of the area in question. As we have already seen, moral realists are willing to admit this point. To stick to Wiggins's formulation, if there were no prospect that arguments in favor of a statement p "should have *non-random efficacy* in securing agreement about whether p", then we would have to conclude that there is simply no fact of the matter (1987b: 149, emphasis added). Realists and their critics part company when it comes to the breadth of convergence that we are supposed to witness in morals. The proponents of the argument from disagreement maintain that there is far more disagreement than agreement on moral matters, while realists deny this. Also, realists insist that the problems apparently posed by the existing amount of irresolvable moral disagreement can be circumvented.

However, one can emphasize that, in the above passage, Williams raises a different problem. Admittedly, Williams's argument can be read along the lines of an argument initially put forward by Gilbert Harman. According to Harman (1977), moral facts are not required to explain (i.e. they do not figure in the best explanation of) our forming moral beliefs. As some authors rightly point out, Harman's argument, usually referred to as "the argument from explanatory impotence" or "the causal argument", is supposed to present us with an independently plausible skeptical challenge. For instance, Nicholas Sturgeon contends that the argument from explanatory impotence is interesting precisely because "there is a *different* difficulty that remains even if we put other arguments for moral skepticism aside and assume, for the sake of argument, that there are moral facts" (1985: 56, emphases in original). The proponents of this argument hold that moral realists fail to establish that moral facts would play a role in the explanation of our moral beliefs. It is worth noting that both the argument from moral disagreement and the argument from explanatory impotence take the form of an inference to the best explanation. However, in the former case the explanandum is the breadth of convergence in moral belief, whereas in the latter case the explanandum is the fact that we hold moral beliefs. It is important to stress that Williams claims that his argument does not turn on whether convergence in moral belief actually occurs. This suggests that Williams's remarks should be interpreted along the lines of the argument from explanatory impotence. According to the proponents of this argument, what matters is whether moral facts figure in the best explanation of our moral beliefs. Assuming that, for instance, realists were able to put forward a convincing argument to the effect that moral facts exert a causal influence on us, the argument from explanatory impotence would be refuted. However, if it turned out that moral facts are explanatorily irrelevant, then the view according to which moral beliefs are the result of a reliable epistemic mechanism would lose its credentials. And the fact that we converge in our moral beliefs could do nothing to change this.

Surely, in order to prove that moral realism is the most viable metaethical position, one would have to show that both the aforementioned skeptical challenges can adequately be met. And if it turns out that the realist cannot provide a convincing answer to these challenges, the moral skeptic can add that the two lines of argument reinforce each other. According to the argument from disagreement, lack of convergence in moral belief is an indicator of the fact that there are no moral facts. Assuming that advocates of the argument from explanatory impotence are right in claiming that moral facts do not figure in the best explanation of our forming moral beliefs, this gives us a further reason for thinking that the link between (lack of) agreement and (non)objectivity is non-accidental. Still, this is not to deny that the two above-mentioned arguments present us with two independently plausible skeptical challenges. The problem raised by the argument from moral disagreement is different from the one raised by the argument from explanatory impotence. The former is the problem of how extensively can the realist use the idea of cognitive errors without undermining his own position. The alleged pervasiveness of moral disagreement would require the realist to postulate widespread cognitive errors. Moreover, the proponents of the argument from disagreement insist that the realist would have to postulate cognitive errors even when it is quite clear that disagreement involves nothing worth regarding as a cognitive failure. So what the realist has to explain is why a significant number of persons, including rational and well-informed persons, fail to get things right. By contrast, the problem raised by the argument from explanatory impotence is the problem of how to account for the fact that we get things right in the moral domain. As it will become apparent in the following chapter, the proponents of this argument hold that moral facts are not required to explain our beliefs about them. Yet surely, if there are moral facts and we can come to know them, then these facts must figure in the best explanation of our forming the moral beliefs that we do. In short, it seems that in order to answer the argument from moral disagreement, the realist has to show that his account of cognitive errors does not end up undermining his own position, whereas in order to refute the argument from explanatory impotence the realist has to establish that moral facts are explanatorily relevant.

All these considerations suggest that, in order to refute the argument from moral disagreement, the realist does not have to deal with the charge of explanatory impotence. This is because the argument from disagreement and the argument from explanatory impotence are supposed to present us with two independent skeptical challenges. However, as it will become apparent in the next section, there are reasons to believe that, in order to conclusively establish that moral realism has no trouble accommodating the distinctive nature of moral disagreement, one would also have to show that moral facts are explanatorily relevant.

#### 2.4 The argument from moral disagreement

Now let me turn to the argument from moral disagreement. According to this argument, the extent and depth of actual moral disagreement undermine the plausibility of moral realism. In response, realists can adopt one of two strategies: (a) question the evidence presented by the proponents of the argument from moral disagreement; or (b) accept this evidence, but argue that it poses no serious threat to moral realism. In what follows, I will consider each of these strategies.

(a) The most common response to the argument from moral disagreement is that, although many cases of moral disagreement appear to resist rational resolution, they are *in principle resolvable*. Some moral realists contend that moral disagreements do not go very deep. For instance, the parties to a debate about the capital punishment may share the view that capital punishment would be morally justified if it were a significant deterrent of crime, but hold different views about whether this is in fact the case. In such a case, moral disagreement can be reduced without remainder to a disagreement over *nonmoral* facts. All that is needed for disagreement to dissolve is reaching a consensus on psychological facts concerning deterrence. Such a consensus may be difficult, if not impossible, to reach in practice. However, just as any other disagreement over nonmoral facts, this is a disagreement that is in principle resolvable.<sup>27</sup> Moral realists commonly emphasize that a great deal of moral disagreement can be traced back to such reasonable, but nevertheless *resolvable*, disagreements over nonmoral facts.<sup>28</sup>

Furthermore, realists argue that some moral disagreements reflect the application of shared moral principles under different circumstances. Appearances notwithstanding, in such cases there is no *genuine* disagreement. For example, the same moral principle may justify certain distributive inequalities in some economic circumstances but not in others. An economically underdeveloped society might promote certain distributive inequalities on the ground that in economically backward conditions distributive inequalities provide incentives that work to everyone's advantage. An economically advanced society might oppose distributive inequalities on the ground that in economically advantage. An economically favorable conditions distributive inequalities are divisive, and therefore, work to everyone's disadvantage. Although at first glance this might look like a case of moral disagreement, disagreement is merely *apparent*.<sup>29</sup>

From a realist perspective, the most natural diagnosis of moral disagreement is that at least one of the parties involved is mistaken. Whereas noncognitivists explain moral disagreement in terms of the parties having conflicting noncognitive mental states, moral realists account for disagreement by reference to the parties' conflicting beliefs. On the latter view, the parties to a moral disagreement assert contradictory moral statements. According to the realist, statements pertaining to moral facts are either true or false. If a

<sup>&</sup>lt;sup>27</sup> As Brink (1984: 118, n14) rightly points out, the argument from disagreement actually relies on the assumption that disagreement over nonmoral facts is always in principle resolvable.

<sup>&</sup>lt;sup>28</sup> Some realists adopt an even stronger position. For instance, Boyd contends that "careful philosophical examination will reveal [...] that agreement on nonmoral issues would eliminate *almost all* disagreement about the sorts of moral issues which arise in ordinary moral practice" (1988: 213, emphasis in original). Other proponents of naturalistic moral realism do not share Boyd's optimism. Compare Railton (1986).

<sup>&</sup>lt;sup>29</sup> For this example, see Brink (1989: 200). For arguments to the same effect, see McGinn (1997: 48) and Sosa (2001: 216). For a discussion of this line of argument, see also Mackie (1977: 37–8).

moral statement is affirmed by one party and denied by the other, it follows that one of the parties is making a false assertion. If so, then one of the parties is subject to some sort of cognitive error. The realist's account of the cognitive shortcomings that explain moral disagreement includes fallacious reasoning, ignorance of the relevant nonmoral facts and impoverished imaginative capacities. Realists rightly stress that imagination is especially important in moral reasoning, given the central role that thought experiments play in the assessment of moral theories.<sup>30</sup> Prejudice, lack of empathy, as well as the emotions and interests that are involved in moral disputes also figure in the realist's account of why people tend to make cognitive errors.

The realist may argue that *any* moral disagreement can be explained by appeal to one of the above-mentioned considerations. According to this line of thought, some cases of moral disagreement are not cases of *genuine* disagreement, while others result from divergent factual beliefs, and therefore, are not cases of genuine *moral* disagreement. As to the cases of genuine moral disagreement, they are always attributable to the parties' cognitive errors (other than the errors resulting from their ignorance of the relevant nonmoral facts). The implicit claim is that any moral disagreement can be settled once the parties have enough time to engage in the debate and no cognitive failure intervenes.

However, this strategy does not take the realist very far. Even if we grant that the realist is right to point out that surface disagreement may be compatible with agreement on deeper moral matters, it is unclear how many cases of moral disagreement can be said to derive from the application of shared moral principles under different circumstances. What is more, it seems that disagreement can persist even when there is a consensus on the relevant nonmoral facts and the parties' beliefs do not result from a failure to reason correctly.

(b) A different strategy is to admit that some cases of moral disagreement are *in principle irresolvable*, but argue that this poses no threat to moral realism. In recent years this strategy has been adopted by an increasing number of moral realists.<sup>31</sup> According to them, some moral issues are indeterminate. Once we accept that morality fails to yield a determinate answer to every practical question, there is no need to explain the failure to converge on a correct answer in terms of the parties' cognitive shortcomings. As Ernest Sosa puts it, if there is no determinate truth as to what ought to be done in certain cases, then "it is understandable that no one should be in a position to mirror [the truth], that no one should enjoy such a competence" (2001: 222).

So far I have been concerned with the version of the argument from disagreement which holds that the extent and depth of *actual* moral disagreement gives us reason to reject moral realism. The appeal to the idea of indeterminacy helps the realist answer this argument. The proponents of this argument contend that the realist's account of moral disagreement involves postulating widespread cognitive errors. However, the argument goes, a more charitable interpretation is that none of the parties to a moral disagreement needs to be mistaken. Furthermore, they claim that the realist has trouble accounting for the fact that moral disagreement seems to persist even among well-informed and rational persons. The suggestion is that the realist has no other option than to postulate cognitive

<sup>&</sup>lt;sup>30</sup> See, for instance, Brink (1989: 202–3).

<sup>&</sup>lt;sup>31</sup> See, for instance, Brink (1987, 202–3). <sup>31</sup> See, for instance, Brink (1984: 116–7; 1989: 202), Boyd (1988: 199–201, 212–4), Wiggins (1990–91), Wolf (1992), Schafer-Landau (1994; 2003: ch. 9), Sturgeon (1994), Thomson (1996: 154, 205–6), and Sosa (2001).

errors even when it is quite clear that disagreement involves nothing worth regarding as a cognitive failure. Yet, by appealing to considerations such as the ones outlined above, the realist offers a promising response to the argument from disagreement. If the realist is right in claiming that some cases of moral disagreement can be reduced to disagreements over nonmoral facts and others can be traced back to divergent circumstances, then this clearly diminishes the force of the argument from disagreement by showing that the role assigned by the realist to cognitive errors in explaining disagreement is less significant than commonly supposed. Moreover, if some moral issues turn out to be indeterminate, we have further reason to believe that the realist is not bound to postulate widespread cognitive errors, and at the same time, we have reason to think that the realist can provide a plausible explanation of why even well-informed and rational persons might fail to converge on the right answer to certain moral questions.

It is important to emphasize that the appeal to the idea of indeterminacy would also enable the realist to answer a different version of the argument from disagreement, which does not rely on any empirical premises.<sup>32</sup> According to this version of the argument, the mere fact that a certain kind of moral disagreement is possible, i.e. disagreement among well-informed and rational persons, gives us reason to reject moral realism. The realist can of course deny that such disagreement is possible, and argue that it is *a priori* that every case of moral disagreement involves a cognitive shortcoming. However, advocates of the above-mentioned argument rightly point out that this move would commit the realist to the existence of a specifically moral cognitive ability. This is because it seems that moral disagreement may occur even when the parties agree on all relevant nonmoral facts and no failure to reason correctly intervenes. The realist can insist that there is some lurking cognitive defect that explains disagreement, but his inability to pinpoint the defect would ultimately lead to positing the existence of a specifically moral cognitive ability.<sup>33</sup> If, on the other hand, the realist admits that there is no lurking cognitive defect by reference to which disagreement can be accounted for, he would have to accept that moral truth is beyond our grasp. Note, however, that by appealing to the idea of moral indeterminacy, the realist can avoid the aforementioned dilemma. As Sosa (2001: 222) rightly emphasizes, once we accept the idea of moral indeterminacy, there is no need to view the parties to a moral disagreement as *differentially competent*. Yet, this is precisely what the proponents of the argument that concerns us here find implausible, i.e. the fact that, on a realist view, some people would have to be considered "morally sighted", while others would have to be considered "morally blind".<sup>34</sup> If, however, a certain moral question admits of no determinate answer, then it is understandable that no one should be in a position to know the answer. Also, it is worth stressing that the realist does not have

<sup>&</sup>lt;sup>32</sup> See Tolhurst (1987) and Wright (1992).

<sup>&</sup>lt;sup>33</sup> There are two things that are worth mentioning here. First, the mere fact that the realist would need to posit the existence of a specifically moral cognitive ability does not mean that he cannot have the upper hand in this debate. Yet, as Tolhurst points out, the realist would have to show that "positing this ability is more than an ad hoc maneuver forced upon [him] by the need to evade a troublesome argument" (1987: 621). Second, it is worth stressing that although moral intuitionism has fallen out of philosophical fashion in the second part of the last century, recent years have witnessed a revival of interest in this view. (See, for instance, Audi (2002), Shafer-Landau (2003), and Zangwill (2006). Even some of the opponents of moral intuitionists urge that this view cannot be dismissed out of hand. See, for instance, Sturgeon (2002).) In what follows, however, I will argue that the realist can avoid positing the existence of a specifically moral cognitive ability. <sup>34</sup> See Tolhurst (1987: 617).

to claim that moral truths are undetectable. It should already be apparent that, on the above-mentioned view, the reason why not even the most competent inquirers can know the right answer to certain moral questions is simply that there is no such answer.<sup>35</sup>

The remainder of this dissertation is devoted to the question of whether the realist can refute the argument from disagreement by appeal to the idea of moral indeterminacy. In this chapter, I will address two difficulties that any proponent of the above-mentioned strategy has to deal with. Before going further, however, a brief remark is in order. In the previous section, I have suggested that even though the argument from disagreement and the argument from explanatory impotence present two independent skeptical challenges, it is doubtful that one can offer a convincing answer to the former argument without also answering the latter one. There are two reasons why this is so.

The first reason has to do with the fact that moral indeterminacy cannot be a pervasive phenomenon. It should be emphasized that those who claim that the argument from disagreement loses its bite once we appeal to the idea of moral indeterminacy explicitly hold that indeterminacy is not pervasive.<sup>36</sup> Surely, to claim otherwise would mean to go against the grain of moral realism. If we found that most moral questions are indeterminate, then we would be justified in concluding that there is no moral reality. Yet, those who propose the above-mentioned reply to the argument from disagreement maintain that moral indeterminacy poses no threat to moral realism provided that there are a significant number of moral questions that admit of determinate answers. Note that the realist is also bound to accept that we can come to know the answers to such questions. (This is because, as we have already seen in the previous section, the realist cannot hold that moral truths are undetectable, nor can he make extensive use of the idea that our cognitive shortcomings prevent us from doing so.) Thus, some might insist that the weight of the realist's argument falls rather on the claim that there are a considerable number of moral questions that collect convergent answers. As already emphasized, realists contend that there is a significant amount of convergence of moral belief. They point out, for instance, that many cases of moral disagreement reflect the application of shared moral principles under different circumstances. Also, they draw attention to the fact that there has been considerable moral progress over time. However, advocates of the argument from disagreement urge, with some plausibility, that disagreement over moral issues is more pervasive than disagreement that occurs in other areas of inquiry. So, the question is whether the realist can win this debate. Some might insist that, at best, we can conclude that there is no decisive evidence one way or the other.<sup>37</sup> Yet surely, if the realist were able to prove not only that there is a significant amount of convergence of moral belief, but also that convergence of moral belief can be regarded as the result of a reliable epistemic mechanism, then this would tip the balance in favor of realism.

There is, however, a second reason why the question of whether moral facts are explanatorily relevant deserves special attention when dealing with the argument from

<sup>&</sup>lt;sup>35</sup> According to Wright (1992), debates about realism and convergence are best interpreted as pivoting on what he calls "Cognitive Command". As he puts it, a discourse exerts Cognitive Command if and only if it is *a priori* that disagreements within the discourse "*unless excusable as a result of vagueness* [...] will involve something which may properly be regarded as a cognitive shortcoming" (1992: 144, emphasis added). Wright suggests that there is a difference in kind between factual and evaluative indeterminacy. For more on this point, see Wright (1994: 334). In what follows, I will call this assumption into question. <sup>36</sup> See, for instance, Wiggins (1990–91: 67, 77, 81) and Sosa (2001: 223).

<sup>&</sup>lt;sup>37</sup> See, for instance, Loeb (1998: 285).

disagreement. If the most plausible response to the argument from disagreement involves defending not just (a) the claim that some moral issues are indeterminate and that our failure to converge on a moral verdict is sometimes due to such indeterminacy, but also (b) the claim that there is a significant amount of convergence of moral belief and that this convergence can be explained on cognitivist assumptions, then a worry might arise about whether an adequate defense of both the aforementioned claims can be mounted from the perspective of the same version of cognitivism. I have emphasized earlier in this chapter that even if moral realism does not have the resources to answer the argument from disagreement, this does not mean that there is no objectivist version of cognitivism that has such resources. More specifically, I have argued that if it can be proved that (nonrelativist) moral constructivism can accommodate the distinctive nature of moral disagreement, this should be considered a satisfactory response to the aforementioned argument. However, in order to establish that the challenge posed by moral disagreement can adequately be met, what needs to be proved is that there is at least one version of cognitivism that has the resources to account for both (a) and (b). And one can argue that defending (a) may be easier if one embraces a version of cognitivism which delivers a weaker form of objectivity, while defending (b) may be easier if a version of cognitivism which delivers a stronger form of objectivity is vindicated. For instance, one can argue that constructivism is better suited to accommodate the idea of moral indeterminacy (since, on this view, moral facts are constructed from the responses of some idealized agents). On the other hand, the argument might go, realism is better suited to prove that convergence of moral belief can be regarded as the result of a reliable belief-producing mechanism (presumably because only realism can accommodate the picture according to which moral beliefs are produced by a causal mechanism<sup>38</sup>). Thus, in order to allay the suspicion that one's reply to the argument from disagreement tacitly trades on intuitions that support different versions of cognitivism, one has to offer a thorough defense of both (a) and (b).

## 2.5 Moral phenomenology

As already pointed out in the Introduction, the view that our failure to reach a moral consensus is sometimes due to there being no determinate answer to the question under consideration is not unproblematic. In the remainder of this chapter, I will briefly address two difficulties that any proponent of the aforementioned view has to deal with.

The first difficulty has to do with the fact that realists claim that moral phenomenology lends support to a cognitivist account of morality.<sup>39</sup> Realists hold that the fact that moral discourse is typically assertoric establishes a *prima facie* presumption in favor of the view that moral sentences are used to express genuine beliefs. Furthermore, they point out that, in both intrapersonal and interpersonal contexts, we act as if there were a right answer to be discovered. As Jonathan Dancy emphasizes, when confronted

<sup>&</sup>lt;sup>38</sup> One might further add that only naturalistic versions of moral realism can accommodate the view that moral facts are causally efficacious. I will discuss this claim in the following chapter.

<sup>&</sup>lt;sup>39</sup> See, for instance, Nagel (1980: 100–1), Dancy (1986), Brink (1989: 23–31), Shafer-Landau (1994: 332, 2003: 23–7), and Smith (1994: 5–6).

with a moral choice, "we struggle to find, not any answer that we can bring ourselves to accept, nor any answer that we can accept in consistency with previous answers, but the right answer" (1986: 172). Otherwise there would be no reason to agonize about what we ought to do. Moreover, realists argue that what lies in the background of persistent moral disagreement is a shared assumption about there being a right answer to the question under consideration. It is precisely because we assume that there is a right answer that it makes sense to we keep on disagreeing. For if we thought that there was no fact of the matter, how could we regard the other party as mistaken? In short, realists stress that various aspects of moral phenomenology are most readily explained on cognitivist (or realist) assumptions.

Yet, if the realist contends that the argument from disagreement loses its bite once we concede that some moral issues are indeterminate, this leads to the following problem. While the realist claims that his metaethical view does justice to moral phenomenology, by arguing that disagreement occurs when a moral question has no determinate answer, he ends up suggesting just the opposite. So, the realist owes us an explanation of how to reconcile the claim that his metaethical view accommodates the phenomenology of moral disagreement with the claim that disagreement arises out of indeterminacy.

Before going further, a brief point of clarification is in order. Some philosophers hold that defending a metaethical view by appeal to commonsense intuitions is dubious. For instance, Richard Joyce (2007) argues that the distinctions involved in defining various metaethical views are often too fine-grained for there to be a clear intuition in favor of one view or another. To illustrate this point, Joyce points out that it is far from obvious whether common sense favors the view that moral facts are mind-independent, given that there are many different ways of articulating the idea of mind-independence. Furthermore, he stresses that the question whether there is a widespread intuition in favor of a certain metaethical view is an empirical question that has not been appropriately investigated. Finally, he points out that even if it turns out that common sense favors a metaethical view, it is not at all obvious how much weight should be assigned to this fact. This is because some metaethical views predict that certain (erroneous) intuitions will be widespread. Therefore, it is doubtful whether the fact that such a metaethical view offends against these widespread intuitions should be allowed to count as a consideration against the view in question.

It should be noted, however, that none of the considerations adduced above is at odds with Joyce's claims. We have seen that realists commonly hold that various aspects of moral phenomenology reflect the cognitivist (or realist) character of commonsense morality. Yet, it is important to keep in mind that, even though realists claim that the way we conduct ourselves in moral inquiry lends support to an objectivist account of morality, they acknowledge that phenomenological considerations are far from being conclusive.<sup>40</sup> Their point is that cognitivism (or realism) offers a natural explanation of certain key aspects of moral phenomenology. So, if it turns out that the most defensible metaethical view is one that offers such a natural explanation, then so much the better. However, realists admit that whether or not we should adopt a certain metaethical view is ultimately to be decided on grounds other than phenomenological ones. More specifically, realists acknowledge that, while various aspects of moral phenomenology aspects of moral phenomenological ones. More specifically, realists acknowledge that, while various aspects of moral phenomenology may reflect our commitment to moral objectivity, they do not necessarily reflect the way things are. In

<sup>&</sup>lt;sup>40</sup> See, for instance, Brink (1989: 24) and Smith (1993: 243-4).

other words, realists admit that commonsense moral thinking might be fundamentally misguided.

Another important thing to note is that some realists emphasize that the view that moral phenomenology reflects our commitment to moral objectivity does not necessarily imply that belief in moral realism is prevalent. For instance, David Brink writes that:

I am willing to admit that, about moral realism, common belief is silent, divided, or even antagonistic. My concern, however, is not with unreflective and untutored metaphysical or metaethical views. My appeal to commonsense moral thinking is not a prediction about the likely results of a Gallup poll on the issue of moral realism. Rather, my concern is with the philosophical implications or presuppositions of moral thought and practice. [...] I claim that cognitivism seems to be presupposed by common normative practices of moral judgment, argument and deliberation [...]. (Brink, 1989: 25)

So, the view that concerns us here does not rely on an empirical claim whose truth has not yet been established. On this view, the *prima facie* advantage that cognitivism (or realism) supposedly enjoys is not established by appeal to considerations about what intuitions people hold.

Finally, let me turn to the suggestion that any attempt to defend a metaethical position by appeal to commonsense intuitions is rendered dubious by the fact that the distinctions involved in defining various metaethical views are often too fine-grained for the folk to have a clear intuition in favor of one view or another. Given that, as I have just emphasized, realists claim that their view derives support from certain deeply entrenched aspects of moral practice rather than from mere intuitions, the aforementioned suggestion seems to lose its relevance. However, there is a point that remains valid even if Joyce is wrong in thinking that the advantage that realism supposedly enjoys lies mainly in the realm of intuition. Let me detail.

Consider a claim that is commonplace among moral realists. According to them, the fact that we see ourselves and others as capable of making moral mistakes reflects our commitment to moral objectivity. Realists point out that, especially when we are faced with a difficult moral choice, we acknowledge that we may not arrive at the right answer. It is precisely because we suppose that there is such a right answer and we fear that we may not arrive at it, the argument goes, that we sometimes agonize about what is the right thing to do. Furthermore, realists emphasize that moral disagreements occur against the assumption that there is a right answer to the question under consideration and that the other party is in error. The fact that we allow for the possibility of moral errors is indicative of the fact that we take the right answer to a moral question to be independent of our beliefs or desires. For if we thought that the criterion for making our moral choices lied within ourselves, there would be no reason to worry that we might make the wrong choice. Likewise, when disagreeing about how to settle a moral question, unless we thought that the right answer was independent of the parties' beliefs or desires, how could we regard the other party as mistaken?

Some realists take a step further and claim that we take the right answer to a moral question to be independent of the attitudes of any *actual* or *hypothetical* moral agent.<sup>41</sup> In other words, they claim that constructivist versions of cognitivism cannot do justice to

30

<sup>&</sup>lt;sup>41</sup> See Brink (1989: 31–2). For a suggestion to the same effect, see Nagel (1980: 100).

moral phenomenology. Yet, one might object to this claim on the same grounds on which Joyce objects to appealing to commonsense intuitions in order to establish that realism enjoys a *prima facie* advantage over other metaethical views. Joyce argues that, even if it turns out that common sense decisively favors the view that moral facts are mind-independent, there are ways of articulating the idea of mind-dependence so as not to conflict with this commonsense intuition. As he rightly points out, while it certainly looks like common sense favors some sort of coarse-grained moral mind-independence, it is far from clear that common sense would oppose the view that whether something is right depends on whether some ideally placed moral agents would approve of it.

In the same vein, one can grant that the way we conduct ourselves in moral inquiry suggests that we hope to arrive at answers whose correctness is independent of our moral beliefs, but argue that our practices do not seem to presuppose the truth of moral realism.<sup>42</sup> Given that philosophers distinguish among subtly different independence relations and that they take such differences to be crucial to the success of one metaethical position or another, establishing that moral phenomenology unequivocally supports a realist account of morality is no easy task. Surely, denying that the correct answer to a moral question is independent of one's beliefs or desires would mean distorting moral phenomenology (as would claiming that the correct answer depends on the practices or conventions that individuals happen to embrace). Nonetheless, one can argue that it is not at all clear that sophisticated versions of (nonrelativist) constructivism are at odds with moral phenomenology. If we identify truth with ideal justifiability, as sophisticated versions of constructivism do, there does not seem to be any difficulty in accommodating the idea that moral mistakes are possible or the idea that the correct answer to a moral question is independent of anyone's beliefs about right or wrong. Admittedly, constructivism would be inconsistent with the phenomenology of moral thought if our common normative practices presupposed that the correct answer to a moral question is independent not only of "here and now" justification, but even of ideal justification. Yet, our practices do not seem to presuppose anything of the sort.

It is important not to lose sight of the fact that what is needed here is not a general argument against identifying truth with ideal justification, but an argument to the effect that our common normative practices presuppose that the two are distinct. According to some realists, pointing out that one can sensibly ask whether ideally justified beliefs are actually true is all that is required to show that our practices presuppose the truth of moral realism. For example, David Brink (1989: 31–4) emphasizes that theories which identify truth with justification cannot represent the skeptical possibility that our moral beliefs, though ideally justified, might nevertheless be false. While realism makes such skeptical worries intelligible, constructivism cannot do justice to them. Therefore, Brink concludes, we should prefer moral realism to moral constructivism.

However, one might find this line of argument unconvincing. As Mark Timmons (1999) rightly points out, there is nothing about our practices that implies that moral error is possible even at the level of ideally justified beliefs. Timmons emphasizes that "even apart from considerations of being ideally situated, it is far from clear that commonsense moral thinking presupposes that moral error is ubiquitously possible" (1999: 91). To illustrate this point, he points out that a moral judgment such as "It is wrong to torture

 $<sup>^{42}</sup>$  For arguing that our common normative practices do not presuppose the truth of moral realism, although they do presuppose the truth of cognitivism, see Copp (1991: 613– 4) and Timmons (1999: 90–2)
innocent children for fun" seems to be immune from error. Furthermore, he stresses that although our common normative practices presuppose that moral error is possible, the subtle difference between the claim that any of our actual moral beliefs might be erroneous and the claim that any moral belief, even ideally justified ones, might be erroneous is not in any way reflected in our practices. So, the realist cannot avail himself of the claim that only his metaethical view can do justice to our ordinary ways of thinking about moral error. Brink's appeal to considerations about the *intelligibility* of raising the question whether even ideally justified beliefs can be false may suggest that his argument to the effect that only realist versions of cognitivism can accurately capture key aspects of moral phenomenology is merely an extension of the argument that moral phenomenology favors cognitivism over noncognitivism. According to the latter argument, cognitivism preserves ordinary talk of moral beliefs being true or false and of people being capable of both moral error and moral improvement. However, in light of the above-mentioned considerations it becomes doubtful that realism is needed to preserve our ordinary ways of thinking about morality. Brink may be right in claiming that we should resist the identification of truth and ideal justifiability, but surely it is not the case that distinguishing between the two would help preserve our ordinary ways of thinking about moral error.

To sum up, it is doubtful that only realist versions of cognitivism can do justice to moral phenomenology. However, I only rely here on the more modest claim that moral phenomenology, and in particular the phenomenology of moral disagreement, creates a presumption in favor of cognitivist theories of an objectivist bent. As already emphasized earlier in this chapter, when discussing the argument from moral disagreement there is no reason to exclude from the outset those versions of moral objectivism that are irrealist.<sup>43</sup> The remarks made above reinforce this point. Insofar as realists are right in suggesting that the argument from moral phenomenology is as an integral part of defending an objectivist position in ethics, objectivist accounts of disagreement that are at odds with moral phenomenology are problematic. Nevertheless, constructivism does not disqualify itself from offering an adequate treatment of disagreement by being inconsistent with phenomenological considerations.<sup>44</sup>

<sup>&</sup>lt;sup>43</sup> In section 2.1 I have also insisted that in order to refute the argument from disagreement it would be enough to show that versions of realism that deliver a weaker form of objectivity (i.e. response-dependent theories) can offer an adequate treatment of moral disagreement. Yet, some authors have argued that response-dependent theories are inconsistent with the phenomenology of moral experience. See Dancy (1986). For a different view, see Smith (1993).

<sup>&</sup>lt;sup>44</sup> I do not mean to imply that whether a particular account of disagreement is plausible ultimately depends on whether it can accommodate the phenomenology. To be sure, there may be reasons for preferring an account of disagreement that would distort the phenomenology. (The argument from disagreement can be understood as offering such a reason. According to the noncognitivist version of the argument, even if our practices seem to presuppose the possibility of error in moral judgment, the problem is that, given the pervasiveness of moral disagreement, cognitivists would have to charge so many of us with error. Thus, although allowing for the aforementioned possibility may initially look like an advantage, it ends up by being an embarrassment for cognitivism.) In response to the argument from disagreement, cognitivists can point out (a) that the extent and apparent intractability of disagreement poses no serious threat to their view, and (b) that their view can easily accommodate the phenomenology of moral disagreement (as well as other aspects of moral phenomenology). If there is good reason to accept (a), then cognitivism can retain the alleged advantage over noncognitivism expressed by (b). Now, surely, if a certain cognitivist view (e.g. nonrelativist constructivism) were inconsistent with phenomenological considerations, this would make it more difficult to offer a convincing account of disagreement. This is so especially since the concern with

Now let me briefly spell out the reasons for doubting that nonobjectivist versions of cognitivism can do justice to moral phenomenology. We have seen that realists typically point out that certain deeply embedded features of moral discourse lend presumptive support to the view that moral sentences express genuine beliefs. Noncognitivists often agree with this claim, i.e. they agree that on this point it is the noncognitivist who bears the burden of proof.<sup>45</sup> Given that according to both error theory and relativism moral sentences are used to express beliefs and are truth-assessable, these views would enjoy the same advantage over noncognitivism. However, both the aforementioned views seem to have difficulties in accommodating other features of moral phenomenology. Consider, first, error theory. As we have already seen, our common normative practices presuppose the possibility of both accuracy and error in moral judgment. Error theory is plainly inconsistent with this aspect of moral phenomenology.<sup>46</sup> Alternatively, one can point out that, even though this view allows for the possibility of error in moral judgment, it does not allow for the kind of error that our practices presuppose possible. I have already emphasized that the fact that we sometimes torment ourselves about the right thing to do, or the fact that we disagree with others about how to settle a moral issue, can only make sense if error in moral judgment is possible. Note, however, that it would not make sense to struggle to find the right answer, nor would it make sense to disagree with others, unless it was possible to be mistaken about specific moral issues. Yet, error theory cannot do justice to this aspect of the phenomenology since, on this view, there being no moral facts guarantees that error in moral judgment is ubiquitous.

Consider now one of the most common criticisms of relativism. According to it, relativism cannot make good sense of moral disagreement.<sup>47</sup> Suppose two persons with different moral outlooks debate whether abortion is morally permissible, one of them holding that abortion is never permissible, the other one claiming that it is sometimes permissible. The relativist would regard these apparently conflicting moral judgments as simultaneously true. On the relativist's view, moral truths are relational, i.e. whether something is morally wrong is relative to one set of standards or another. When a person judges that abortion is never permissible, the relativist story goes, this judgment is made in accordance with the moral standards she accepts. The same holds for the person who judges that abortion is sometimes permissible. If each of these moral judgments relates to

<sup>46</sup> See, for instance, Timmons (1999: 77–80).

whether one's metaethical view can accommodate various features of moral phenomenology is also shared by some noncognitivists. (See, for instance, Blackburn (1993). See also Stevenson (1948) for the claim that the phenomenology of moral disagreement is best explained on noncognitivist assumptions.)

<sup>&</sup>lt;sup>45</sup> Contemporary noncognitivists attempt to prove that the realist-seeming grammar of moral discourse can be explained on noncognitivist assumptions. (See Blackburn (1993)). However, I will not address here the question whether such attempts are successful. This is because even if it turns out that noncognitivism can make sense of various aspects of moral phenomenology, it remains true that moral objectivism offers the most *natural* explanation of the phenomenology. (The same point applies to relativism. Even if turned out that some sophisticated version of relativism can accommodate the phenomenology, we would, other things being equal, prefer the more natural explanation.) Also, note that phenomenological considerations are not offered here as an argument in favor of moral objectivism. My intention is simply to draw attention to the fact that, since objectivists often make use of the argument from phenomenology, it is incumbent on them to offer an account of disagreement that is not inconsistent with phenomenological considerations.

<sup>&</sup>lt;sup>47</sup> See, for instance, Lyons (1976: 119–20), Sturgeon (1994), Timmons (1999: 126–7), and Horgan and Timmons (2006: 228–9).

a different set of standards, then ascribing truth to both of them seems to avoid endorsing inconsistencies.

The problem is that relativism is seriously at odds with moral phenomenology. The relativist is committed to regarding the parties to a disagreement of the kind described above as merely talking past each other. If moral language is used as the relativist supposes, i.e. if the use of moral terms is governed by the moral standards the speaker accepts, it follows that two persons who debate whether abortion is morally permissible, but who do not share the same moral outlook, would in fact use moral terms to refer to different properties. This means that there can be no genuine disagreement between such persons. Yet surely, we conceive of the parties to a dispute about the permissibility of abortion as disagreeing (rather than merely talking past each other) even if they do not share the same moral outlook. It is worth noting that relativism would not only distort the phenomenology of moral disagreement, but would also implausibly suggest that it is not possible for two persons with different moral outlooks to agree about whether abortion is morally permissible. Any such agreement would be illusory since statements to the effect that abortion is morally wrong or that it is not morally wrong must be understood as referring to different properties. Another important aspect of moral phenomenology that is incompatible with relativism is the following. As Nicholas Sturgeon (1994: 100, 112) rightly emphasizes, when other competent persons who embrace values different from our own disagree with us on a particular moral issue, we take this fact to provide a challenge to our own views. However, if the relativist is right in thinking that such disagreements are not genuine, then they cannot provide any reason for doubting our own views.

This concludes my discussion of nonobjectivist versions of cognitivism. I turn now to the question whether objectivist versions of cognitivism succeed in offering an account of disagreement that is consistent with phenomenological considerations. We have seen that realists often claim that moral phenomenology, and in particular the phenomenology of moral disagreement, lends support to an objectivist account of morality. In other words, they claim that objectivist metaethical views enjoy a *prima facie* advantage over nonobjectivist ones because only the former can do justice to moral phenomenology. This being the case, if the objectivist's response to the argument from disagreement required construing disagreement in a way that was inconsistent with the phenomenology, then, by the objectivist's own light, this would not be an adequate response to the challenge raised by the nonobjectivist.

Recall that realists typically point out that moral disagreements occur against the background assumption that there is a right answer to the question under consideration. So, the realist owes us at least an explanation of how to reconcile the claim that realism does justice to moral phenomenology with the claim that disagreement arises out of indeterminacy.

Three lines of response suggest themselves. First, those who appeal to the idea of moral indeterminacy in answering the argument from disagreement can emphasize that indeterminacy is not a pervasive phenomenon and that most moral questions admit of right answers. The idea is that the account of disagreement that they propose construes most cases of disagreement as cases in which there is a right answer to be discovered, and therefore, it is not at odds with the phenomenology of moral thought. If the realist adopts this line of response, then there is a further reason for thinking that an adequate reply to the argument from disagreement cannot exclusively rest on the idea that some moral issues are indeterminate. In order to mount a convincing reply to this argument, the realist would also have to prove that lack of consensus on a whole range of moral issues that are not indeterminate poses no serious threat to realism. A combined strategy like the one outlined in the previous section can achieve this end, or so I will argue.

Second, those who defend the account of disagreement that concerns us here can point out that even when there is no uniquely right answer to the question of what morally ought to be done, there may be a bunch of wrong answers.<sup>48</sup> This point may be obscured by the fact that many of the questions used to illustrate the pervasiveness and apparent intractability of moral disagreement are framed, and sometimes rightly so, as "yes or no" questions (e.g. "Is abortion morally permissible?", "Is it morally right to raise animals for food?"). However, not all questions that elicit disagreement are so framed. Sometimes the question in need of an answer is "What is the morally right thing to do under the circumstances?" In such cases, even if morality does not yield a univocal verdict about what ought to be done, there may still be many courses of action that are plainly wrong. To illustrate this point, consider Jean-Paul Sartre's (1956) famous example involving a young man torn between staying home to take care of his old mother and joining the Free French to fight against Nazi Germany. Now, supposing that the most defensible moral theory fails to yield a uniquely correct answer about the right thing to do in such circumstances, surely this does not mean that there can be no determinately wrong courses of action. For one could plausibly argue that it would have been wrong of the young man to choose not to discharge any of aforementioned duties, i.e. the duty to assist his old mother and the duty to his country in times of war. In short, the view that there are indeterminate moral issues would accurately reflect a key aspect of the phenomenology of moral thought. Since even in cases of indeterminacy there may be morally wrong choices, there is no reason to doubt that *anything real is at stake* in such cases.

A third line of response is suggested by a remark made by Sturgeon. Recall that, according to Sturgeon, relativism is objectionable because it cannot accommodate the fact that we take disagreement with persons who seem competent in ways relevant to the issue at stake to be a direct challenge to our own views. That we do so is precisely because we care about whether our moral views are correct and, according to our common understanding of disagreement, the parties cannot both be correct. The relativist view that disagreements are not genuine and that the parties are merely talking past each other would strip us of any motive for trying to understand the other party's point of view and for thinking that we may learn from doing so. Yet, as Sturgeon (1994: 113) notices, the same objection would apply to those versions of realism that allow for indeterminacy. For one can argue that there is no reason to pursue a question that admits of no uniquely right answer.

However, Sturgeon rightly points out that:

[A] nihilist view of an issue we disagree about *need not* discourage discussion and mutual learning at all. [...] Even if my friend and I suspect that our question as initially framed has no uniquely correct answer, we learn in discussing it. One of

<sup>&</sup>lt;sup>48</sup> For arguing that the view according to which some moral issues are indeterminate does not imply that "anything goes", see Wolf (1992: 790).

the most important things that we learn is how to reframe the question, and in my example we learn that partly from one another. [...] This seems to me to happen all the time with more complex questions in almost all disciplines. We debate and discuss disagreements, fully anticipating that one result will be refinement or replacement of the initial question, in the light of distinctions the disagreeing parties force on one another's attention. (Sturgeon, 1994: 113, emphasis in original)

In other words, disagreement remains a stimulus to inquiry even when the moral issue we disagree about is indeterminate.

Yet, one obvious difficulty for the view that moral disagreement is sometimes due to indeterminacy has not been addressed so far. Recall that the idea that there are indeterminate moral issues is supposed to account for the possibility of disagreement among agents who suffer from no cognitive shortcomings. This reply to the argument from disagreement runs into the following difficulty. Presumably, agents who suffer from no cognitive shortcomings would have to be aware of all morally relevant facts, including those which can render a situation morally indeterminate. Therefore, it seems doubtful that such agents would continue to disagree about how to settle a moral issue instead of just agreeing on its indeterminate character. This is probably the most common and the most powerful objection to the view that some cases of moral disagreement can be traced back to there being no determinate answer to the question under consideration. However, given that an investigation of the sources of moral indeterminacy may shed some light on how to answer this objection, I will postpone its discussion until the final chapter.

One suggestion I will consider is that the crux of the matter is not whether moral objectivists can account for *disagreement* among agents who suffer from no cognitive shortcomings, but rather whether they can account for *lack of convergence* on a single moral verdict among such agents. (Note that, in claiming that even when indeterminacy becomes a matter of shared awareness, it still makes sense to keep on discussing the issue at stake, Sturgeon is in fact adopting a similar line of thought. According to him, the process of reasoning with one another is supposed to eliminate the cognitive errors of the parties (1994: 105, 113). In other words, if the parties to a moral disagreement suffer from no cognitive shortcoming, they will soon come to realize that there is a sense in which each of the parties involved was right. This realization brings with it a further realization, i.e. that while each party was right about some subsidiary issue, they were both wrong about there being a determinate answer to the question under consideration. Yet, as Sturgeon rightly points out, this realization need not signify that discussing further the issue at stake is pointless.) This line of argument allows the objectivist to stay true to the view that, absent a belief in there being a uniquely right answer to the question under consideration, it does not make sense to *disagree* (i.e. to assert one's point of view over and against the opponent's point of view). Still, the question is whether, by shifting the focus from whether there can be *disagreement* among agents who suffer from no cognitive shortcoming to whether such agents may sometimes *fail to converge* on a single moral verdict, the objectivist can meet the challenge posed by the argument from disagreement.

#### 2.6 Realism and indeterminacy

# 2.6.1 Dummett on realism and indeterminacy

In what follows, I will address another worry about the view that lack of consensus on some moral issues is due to indeterminacy, namely that realism and indeterminacy are incompatible. The assumption that realism about an area of discourse would require an unqualified endorsement of the principle of bivalence (i.e. the principle that every statement in the discourse is determinately either true or false) is often ascribed to realists by their critics. However, some realists argue, persuasively in my view, that abandoning the principle of bivalence for some statements does not imply a form of antirealism.

Let me first try to spell out the reasons for thinking that realism and indeterminacy are incompatible. The most prominent advocate of the idea that realism involves a commitment to the unrestricted principle of bivalence is Michael Dummett. So, it is perhaps best to start by briefly considering his characterization of realism. Dummett was primarily concerned with putting forward a characterization of realism that enabled us to identify a unifying thread among apparently disparate philosophical disputes in which the term "realism" figures. According to him, what is mainly at issue in disputes among realists and anti-realists is whether statements in a given class have evidencetranscendent truth-conditions. While Dummett takes the view that truth is evidentially unconstrained to be the hallmark of realism, he is equally willing to characterize realism in terms of a commitment to the unrestricted principle of bivalence. Consider a much quoted passage from Dummett:

The very minimum that realism can be held to involve is that statements in the given class relate to some reality that exists independently of our knowledge of it, in such a way that reality renders each statement in the class determinately true or false, again independently of whether we know, or are even able to discover, its truth-value. (Dummett, 1982: 55)

There are two distinct ideas in this passage. First, there is the idea that a statement in the disputed class may be true or false independently not only of whether we actually come to know its truth-value, but also of whether it is, even in principle, possible for us to know its truth-value. Second, there is the idea that every statement in the disputed class is determinately either true or false. It is worth noticing, however, that the relationship between these two ideas is less straightforward than it may at first seem. Surely, it is tempting to think that commitment to the unrestricted principle of bivalence is a sufficient condition for realism. This is because claiming that every statement in the disputed class, including statements which are undecidable, must be determinately true or false would involve thinking that such statements are capable of truth or falsity in the absence of evidence one way or the other. In other words, claiming that the principle of bivalence that statements in that class have evidence-transcendent truth-conditions.<sup>49</sup>

<sup>&</sup>lt;sup>49</sup> Some commentators have raised doubts about whether endorsing the principle of bivalence is sufficient to yield realism. For instance, Wright (2003: 14) argues that someone who claims that statements about the mental can be translated without loss into statements about behavior may consistently endorse the principle of bivalence. However, the argument goes, such a thoroughgoing reductionist view is hardly worth calling

However, it is far from obvious that commitment to the unrestricted principle of bivalence should be viewed as necessary for realism. As many critics have rightly pointed out, refusal to assert unrestrictedly the principle of bivalence could be based upon reasons that are compatible with the view that truth is evidentially unconstrained.<sup>50</sup> For instance, it is open to the realist to hold that statements containing empty singular terms (such as "The present king of France is bald") are neither true nor false.<sup>51</sup> Moreover, and more importantly for the purposes of the argument developed in this section, the realist can maintain that some statements containing vague predicates do not have a determinate truth-value. More will be said later on about whether acknowledging this possibility is consistent with realism. For now, let us simply note that holding that bivalence fails for certain statements within the disputed class need not signify a departure from the view that truth is evidentially or epistemically unconstrained. Commitment to the abovementioned view crucially involves the belief that, when a statement has a determinate truth-value, what makes the statement in question true or false is independent of our knowledge, or even of our capacity for gaining such knowledge. Arguably, retaining the belief that truth is not essentially constrained by our epistemic capacities and, at the same time, holding that vague statements do not possess a determinate truth-value is a coherent possibility.<sup>52</sup>

To recap, refusal to assert bivalence does not always signal the adoption of the view that truth is essentially evidentially constrained. In other words, we have reason to reject Dummett's suggestion according to which refusal to assert the principle of bivalence is

realism about the mental. Two points are worth emphasizing in this connection. First, although in some passages Dummett claims that the principle of bivalence is the touchstone of realism, in other passages he explicitly denies that endorsing bivalence is a sufficient condition for realism. See, for instance, Dummett (1978*a*: xxix). Compare Dummett (1993: 467). Second, Dummett (1978*a*: xxxi) explicitly states that his characterization of realism does not capture what is at issue in *all* disputes among realists and their opponents. He claims that even though the term "realism" has been traditionally used to cover, *inter alia*, the issue of reductionism, unless this issue is sharply distinguished from the issue of whether statements of a given class can be held to be determinately true or false, no progress can be made in discussing either of them. So, given that Dummett sets aside the issue of reductionism in characterizing realism, the claim that endorsing bivalence is the criterion for being a realist should be taken to mean that this is the criterion for embracing the view that truth is evidentially unconstrained. Obviously, this does not mean that one cannot object to Dummett's characterization of realism on the grounds that it fails to capture what is most deeply at issue in disputes among realists and their opponents. However, the question whether endorsing bivalence is sufficient for realism in Dummett's sense should be distinguished from the question whether endorsing bivalence is sufficient for whatever conception of realism is taken to be most appropriate.

<sup>&</sup>lt;sup>50</sup> See, for instance, Rosen (1995), Hale (1997: 274), McGinn (1999), and Wright (2003: 14).

<sup>&</sup>lt;sup>51</sup> According to Dummett (1981: 483), the view that bivalence does not hold because some singular terms fail to refer involves a kind of anti-realism (i.e. it involves rejecting an extreme realist position which inflates ontology with Meinongian objects). However, in one of his later works, Dummett writes that he has struggled "to find a principled distinction between deep and shallow grounds for rejecting bivalence, the latter being compatible with realism and exemplified by the truth-value gap recognized, in different ways, by Frege and by Strawson, as induced by empty subject-terms" (1993: 467).

<sup>&</sup>lt;sup>52</sup> As an illustration of this point, consider statements about the past. Intuitively, claiming that the truthvalue of such statements is independent of whether we can have evidence for them is quite different from claiming that *all* statements in this class have a determinate truth-value. More to the point, it is difficult to see why someone who held that statements about the past have evidence-transcendent truth conditions would thereby be committed to thinking that a statement such as "John was bald", said of a man about whom there is evidence that he had several hundred hairs scattered over his scalp, *must* be determinately either true or false.

invariably prompted by anti-realist beliefs. It should be noted, however, that Dummett's characterization of realism differs from the characterization of realism put forward here. Recall that at the beginning of this chapter I have characterized moral realism in terms of mind-independence. So, the question that I need to tackle is whether there is any reason to believe that realism in this sense is incompatible with indeterminacy. Yet, given that Dummett contends that framing the issue between realists and anti-realists in semantic terms is preferable to framing the issue in the traditional metaphysical jargon, let me first clarify my reasons for sticking with the traditional way of depicting the issue between realists and their opponents.

Dummett holds that disputes between realists and anti-realists are best understood as relating to a class of *statements*, rather than to a class of entities. More concretely, he argues that the disagreement between realists and their opponents is a disagreement over the kind of *meaning* that statements of the disputed class possess. According to Dummett, the realist holds that "the meanings of statements of the disputed class are not directly tied to the kind of evidence for them that we can have, but consist in the manner of their determination as true or false by states of affairs whose existence is not dependent on our possession of evidence for them" (1978*b*: 146). By contrast, the anti-realist insists that "the meanings of these statements are tied directly to what we count as evidence for them, in such a way that a statement of the disputed class, if true at all, can be true only in virtue of something of which we could know and which we should count as evidence for the disputed class the main point of contention between realists and anti-realists.

As already emphasized, Dummett's aim is to offer a characterization of realism that would enable us to identify a unifying thread among apparently disparate disputes in which the term "realism" figures. The question is whether Dummett's characterization of realism succeeds in capturing what is mainly at issue in disputes among realists and antirealists. There are several reasons for doubting that this is the case. First, one can argue that Dummett was wrong in thinking that "the theory of meaning underlies metaphysics" (1978a: xi). More specifically, one can point out that some disputes between realists and their critics do not seem to concern the notion of truth that is appropriate for statements of a given class.<sup>53</sup> The dispute between moral realists and error theorists is a case in point. As we have already seen, if second order ethics were confined to analyzing the meaning of moral statements, error theorists would have no quarrel with realists. The disagreement between realists and error theorists is a disagreement over whether there are moral facts in virtue of which some moral statements can turn out to be true. Dummett's way of framing the issue between realists and their opponents cannot properly account for the above-mentioned dispute. In fact, according to Dummett's characterization of realism, error theorists would count as realists, since they have no complaint against the realist's account of the truth conditions of moral statements.

Second, one can argue that, even if most disputes between realists and their critics can be construed as being concerned with the notion of truth appropriate for statements of the disputed class, it is far from obvious that it is always the notion of potentially evidence-transcendent truth that is at stake in such disputes. For instance, Crispin Wright points out that this way of construing the debate precludes realism about statements for

<sup>&</sup>lt;sup>53</sup> For this line of argument, see Hale (1997: 286–7). For expressing doubts about whether linguistic issues can successfully be substituted for metaphysical issues, see also McGinn (1999).

which evidence, favorable or unfavorable, is guaranteed to be available. Wright argues that "it ought to be possible to take a realist view of what makes for the truth or falsity of statements whose truth values are not conceived as evidence-transcendent" (2003: 15). However, the objection goes, Dummett's anti-realist would have no reason to oppose realism in the aforementioned sense.

# 2.6.2 Comparative indeterminacy

In what follows, I will leave aside Dummett's characterization of realism and focus on the question whether there is good reason to believe that realism as characterized here, i.e. in terms of mind-independence, is incompatible with indeterminacy. Given that my interest is in *moral* realism, and that the criticisms of Dummett's characterization of realism raise doubts about whether it is even possible to come up with a characterization of realism that would serve to distinguish all species of realism from the corresponding species of anti-realism, irrespective of the area of discourse, henceforth my discussion will mainly focus on moral realism.

As I have already pointed out at the beginning of this chapter, moral realism is commonly taken to involve the claim that there are moral facts and that the existence and nature of these facts is mind-independent. Since there are different ways of articulating the idea of mind-independence, this characterization of moral realism is not wholly unproblematic. One way of articulating the idea that moral facts are mind-independent is to say that the existence and nature of these facts is independent of the attitudes of any actual or hypothetical moral agent. This claim might be taken to imply that moral facts would exist even if there were no moral agents. However, this is not what moral realists have in mind, since it is hard to see how there could be any instances of right or wrong in the absence of moral agents.<sup>54</sup> What moral realists mean is rather that moral facts are existentially and conceptually independent of the attitudes of any particular moral agent. Also, they insist that moral facts are not a function of the attitudes of some hypothetical moral agent(s).

What reason is there to think that realism as characterized above is incompatible with indeterminacy? Whatever else they disagree about, moral realists share the view that the existence and nature of moral facts does not depend on anyone's attitudes.<sup>55</sup> To put it another way, according to moral realists, the instantiation of moral properties is not determined by our beliefs or by classifications of our own making.<sup>56</sup> However, if realists

<sup>&</sup>lt;sup>54</sup> See, for instance, Shafer-Landau (2003: 15) and Cuneo (2007a: 46-7).

<sup>&</sup>lt;sup>55</sup> One point of contention among realists is whether moral facts exist independently not only of whether we actually come to know them, but also of whether it is, even in principle, possible for us to gain knowledge of them. For instance, Shafer-Landau (2003: 17) holds that epistemically ideal agents at the Peircean limit of inquiry must be aware of all moral facts. Compare Brink (1989: 15ff). Further disagreement concerns the issue of whether moral facts exist entirely independent of characteristic human responses. Given that response-dependent theories may introduce new sources of indeterminacy into the picture, the question that arises is to what extent response-dependence compromises realism. I will come back to this question later on.

<sup>&</sup>lt;sup>56</sup> Talk of moral facts and talk of moral properties can be seen as largely intertranslatable. One can say, for instance, either that a certain action has the property of being morally wrong, or that it is a moral fact that the action in question is morally wrong. As Shafer-Landau puts it, "moral facts register the instantiation of

are right in claiming that there are moral properties "out there", then how can there not be a determinate answer to the question whether a particular moral property is instantiated in a given case?<sup>57</sup>

If the above formulation accurately captures the intuition that speaks against the compatibility between realism and indeterminacy, then showing that moral realism can accommodate the idea that some moral issues are indeterminate may be an easier task than is commonly supposed. Let me detail.

Following Shafer-Landau (1995), I will henceforth distinguish between *alethic indeterminacy* and *comparative indeterminacy*. According to the definition proposed by Shafer-Landau, we are confronted with a case of alethic indeterminacy if there is no determinate answer to the question whether a particular moral property is instantiated by a given person or action. By contrast, comparative indeterminacy arises when there is no right answer to the question of which of two persons or actions better exemplifies a particular moral property.<sup>58</sup>

Let me start by focusing on comparative indeterminacy. When we engage in moral reasoning, we are often interested in which of two persons or actions better exemplifies a particular moral property. We may want to know, for instance, which of two persons is more tolerant, or which of two actions is more just. Most often, making such comparisons does not raise any special difficulties. In some cases, however, we seem to be unable to reach any of the verdicts (a) a is more F than b, (b) a is less F than b, or (c) a and b are equally F (where a and b stand for two objects of moral assessment that exemplify the moral property F-ness). The question is whether this type of indeterminacy is inimical to moral realism. Obviously, if what is puzzling about the idea that moral realism can allow for indeterminacy is the possibility of there being no determinate answer to the question

moral properties" (2003: 65). Note, however, that the notion of a moral fact is somewhat broader than that of a moral property. For example, someone who countenances the existence of moral properties may want to claim that it is a moral fact that one (morally) ought to take a certain course of action. A further complication has to do with the fact that, consistently with viewing moral facts as the truth-makers of true moral statements, one can refuse to countenance the existence of moral properties. Although this position is typically adopted by constructivists, this is not always so. One can, for instance, hold that statements about what we have reason to do can be true or false independently of our beliefs or inclinations, while rejecting the idea that there are moral properties. (Such a view is advocated by Nagel (1980; 1986). Nagel insists that there is no "extra set of properties of things and events in the world" (1986: 140)). However, in what follows I will ignore these complications.

<sup>&</sup>lt;sup>57</sup> For this way of articulating the intuition against the compatibility between realism and indeterminacy, see Shafer-Landau (1995: 84).

<sup>&</sup>lt;sup>58</sup> Shafer-Landau also refers to alethic indeterminacy as metaphysical indeterminacy. He emphasizes that alethic indeterminacy (i.e. truth-value indeterminacy) arises in cases of noncomparative moral assessment, i.e. when we are interested simply in whether a person or an action exemplifies a particular moral property. Shafer-Landau (1995: 85) contends that the sense in which there is no right answer in cases of comparative moral assessment is different from that involved in noncomparative cases, and by that he means both that comparative indeterminacy does not require abandoning the principle of bivalence. However, more argument is needed to establish the latter claim. According to Shafer-Landau, comparative indeterminacy arises when the trichotomy of relations *better than, worse than*, and *equally good* fails determinately. It is precisely because he assumes that in comparative cases the truth-values of all the relevant statements are determinately false that he concludes that the principle of bivalence would hold good. Yet, as it will become clear in Chapter 4, the question whether the failure of the standard trichotomy of value relations must be understood as a *determinate* failure or rather as an *indeterminate* one is the subject of philosophical controversy.

of whether a particular moral property is instantiated in a given case, then comparative indeterminacy poses no threat to moral realism. This is because the aforementioned type of indeterminacy is not due to qualms about whether certain objects of moral assessment would instantiate a given moral property F-ness, but to there being no right answer to the question of how two objects of assessment a and b compare with respect to F-ness.

Still, one might attempt to resist the view that moral realism has no difficulty in accommodating comparative indeterminacy on the following grounds. One might argue that the lack of a right answer to questions of comparative moral assessment is inimical to realism not because it points to an indeterminacy about whether a particular property is instantiated in a given case, but because it threatens to lead to inconsistencies. Let me explain. Suppose we are confronted with a choice between two incompatible courses of action A and B, and that deciding between them would involve comparing A and B with respect to F-ness (e.g. justice). Suppose further that, after careful deliberation, we reach the conclusion that A is neither more nor less F than B, nor are they equally F. Since what is supposed to guide our choice in this example is the degree to which the available courses of action instantiate the moral property of F-ness, it seems that we are morally required to do both A and B. However, we have assumed from the outset that A and B are incompatible courses of action, so that doing B involves not doing A, and vice versa. So, in this example we would be morally required both to do A and not to do A. Thus, allowing for comparative indeterminacy seems to lead to formal inconsistency.

However, one can argue that, even though statements asserting that one would be morally required both to do A and not to do A must be seen as conflicting, there is no reason to think that such a conflict entails a logical inconsistency. If such statements were proven to be contradictory, that would be bad news for moral realism. Yet, as numerous authors rightly point out, conflicting action-guiding statements may be simultaneously true. For instance, Philippa Foot emphasizes that consistency among such statements is "easily explicable on a 'because of this ..., but because of that ...' basis" (1983: 391).<sup>59</sup> According to this suggestion, the fact that an act A would have the natural property x is a reason for doing A, whereas the fact that an incompatible act B would have the natural property y is a reason for doing B. Since doing A would involve not doing B, it follows that someone trying to decide between these courses of action would have reasons to do A and, at the same time, there are reasons not to do A involves no contradiction.

One can, however, raise doubts about whether the charge of inconsistency can be avoided by characterizing the type of moral conflict that occurs in cases of comparative indeterminacy in terms of conflicting moral reasons. More specifically, one can contend that such a characterization is unsuitably weak for the aforementioned cases. We are well accustomed with cases in which there are reasons both *for* and *against* adopting a particular course of action. The presence of such reasons is certainly not peculiar to cases of comparative indeterminacy, but is rather a common feature of our moral life. However, in most cases, the reasons that speak in favor of adopting a course of action override, or are overridden by, the reasons that speak against it. By contrast, in cases of comparative indeterminacy, the reasons on one side are not overridden by reasons on the other side, and this leaves the agent at a loss about what morally ought to do be done. Thus, one can argue that the later type of moral conflict should be characterized in terms of conflicting

CEU eTD Collection

<sup>&</sup>lt;sup>59</sup> For suggestions to the same effect, see Zangwill (1999) and Schaber (2004).

*moral oughts* rather than in terms of conflicting *moral reasons*.<sup>60</sup> Yet surely, the argument would go, claiming that one both ought and ought not to adopt a course of action involves some sort of inconsistency.

The realist who wants to allow for indeterminacy can respond that the appearance of inconsistency arises only on a particular interpretation of the oughts that are in place in cases of comparative indeterminacy. If the conflicting oughts one is confronted with in such cases were all-things-considered oughts, then it would follow that such oughts vield a contradiction. However, this interpretation is not forced upon us as the most adequate interpretation of the oughts that arise in the above-mentioned cases. This point is often emphasized in the literature on moral dilemmas. Moral dilemmas are commonly defined as situations in which an agent ought to do two acts, can do each, but cannot do both. Opposition to moral dilemmas takes several forms, one of which is holding precisely that the existence of such cases would point to some sort of inconsistency.<sup>61</sup> A common line of defense against this charge is to claim that moral dilemmas are not to be conceived of as conflicts between all-things-considered moral oughts, but merely as conflicts between competing non-overridden moral oughts.<sup>62</sup> Although some might still find troublesome the idea that morality sometimes fail to be uniquely action-guiding, once it is made clear that there can be no conflicts between all-things-considered moral oughts, the claim that moral dilemmas reveal an inconsistency will have to be abandoned as untenable.

At this point, one might raise the following objection. One might point out that the notion of a moral dilemma is a philosophical term of art and that there is considerable controversy over how to define moral dilemmas. Yet, given that numerous philosophers have denied that moral dilemmas are possible, no definition can fulfill its purpose unless it captures what is taken to be controversial about their possibility. Thus, one might argue that, if there is to be any substantial disagreement between proponents and opponents of moral dilemmas, dilemmas must be defined as cases in which there is a conflict between

<sup>&</sup>lt;sup>60</sup> A further assumption may be needed here, namely that the relevant reasons on each side are sufficiently strong to yield an obligation. Consider an example in which one has trouble deciding to which of two charitable foundations to give money away. Though it may be true that one has moral reasons to donate money to each of the two foundations, it may plausibly be argued that one is *not* morally obliged to donate money to either of them. This is because, although one may have a moral obligation to contribute to charity, there is a considerable leeway as to the time and the way in which this obligation is to be discharged. So, not all cases of comparative indeterminacy can be characterized in terms of conflicting moral oughts.

<sup>&</sup>lt;sup>61</sup> The existence of moral dilemmas may also seem problematic given that moral dilemmas combined with certain widely held principles of deontic logic (e.g. the agglomeration principle or the "ought" implies "can" principle) lead to paradoxical implications.

<sup>&</sup>lt;sup>62</sup> This suggestion is present in the writings of most advocates of moral dilemmas, although the terminology used to convey it differs among them. For instance, Foot (1983) distinguishes between two types of ought statements, and argues that only "type 2 ought statements" are asserted on an all things considered basis; Dancy (1993: 110–1) distinguishes between "comparative" and "non-comparative practical oughts"; Zangwill (1999) distinguishes between "*pro tanto* obligations" and "overall" or "all-things-considered obligations". See also Guttenplan (1979–80: 64, 75) for the claim that the word "ought" may be too strong, and thus unsuited to capture the conflicting moral demands that are present in a moral dilemma. Sinnott-Armstrong (1988: 17–21) argues against construing moral dilemmas in terms of a conflict between "strict", "absolute" or "overriding" moral requirements. Finally, Schaber (2004) explicitly denies that moral dilemmas are to be defined in terms of "conclusive" or "all-things-considered oughts".

two all-things-considered moral oughts.<sup>63</sup> Presumably, this is because opponents of moral dilemmas allow that there are cases when equally stringent moral considerations demand incompatible acts, but insist that in such cases one has an all-things-considered obligation to perform either one act or the other.<sup>64</sup>

In what follows, I will offer some brief suggestions to the effect that this way of construing the debate between friends and foes of moral dilemmas is misguided. More specifically. I will point out that, even if we do not attribute to friends of dilemmas the view that conflicts between all-things-considered moral oughts are possible, there is no reason to doubt that there is anything substantial at issue in this debate.

Admittedly, a main point of contention between proponents and opponents of moral dilemmas is whether moral conflicts are soluble without remainder. In his influential paper "Ethical Consistency", Bernard Williams wrote that:

It seems to me a fundamental criticism of many ethical theories that their accounts of moral conflict and its resolution do not do justice to the facts of regret and related considerations: basically because they eliminate from the scene the *ought* that is not acted upon. [...] Such an approach must be inherent in purely cognitive accounts of the matter; since it is just a question of which of the conflicting *ought* statements is true, and they cannot both be true, to decide correctly for one of them must be to be rid of error with respect to the other [...]. (Williams, 1973a: 175, emphases in original)

Even though Williams was not mainly interested in tragic conflicts, he believed that such extreme cases shed some light on the structure of moral conflict. According to him, the fact that in such cases the agent feels regret no matter what course of action he chooses to adopt is indicative of the fact that the agent thinks that he has failed to do something that he ought to have done, even while thinking that he has adopted the right course of action. Moreover, our viewing regret as morally appropriate in such cases suggests that the agent is right in thinking that the moral ought that was not acted upon is not annulled. Williams thought that cognitivism cannot adequately represent the structure of such conflicts. This is because he believed that, on a cognitivist picture, conflicting ought statements cannot simultaneously be true. As he put it, the cognitivist is committed to viewing the structure of moral conflict as similar to the structure of belief conflict, and thus, to holding that one of two conflicting ought statements must be rejected. However, the structure of moral

<sup>&</sup>lt;sup>63</sup> For instance, see Connee (1989: 134), for arguing against Sinnott-Armstrong that a conflict between nonoverridden moral requirements "could not constitute the philosophically controversial sort of dilemma". See also Brink (1996), for claiming that the type of moral conflict that has paradoxical implications is a conflict between all-things-considered moral oughts and, therefore, it is for this type of conflict that we should reserve the term "moral dilemma". This is a rather strange line of argument, since advocates of dilemmas argue precisely that moral dilemmas do not have paradoxical implications. Brink, like many other opponents of dilemmas, distinguishes between prima facie and all-things-considered oughts and assumes that, when characterizing the conflicting moral demands that are in place in a dilemmatic situation, one must opt between the two aforementioned descriptions. Thus, if moral dilemmas are characterized in terms of conflicting prima facie oughts, then there seems to be no substantial disagreement between friends and foes of dilemmas. If, on the other hand, moral dilemmas are characterized in terms of conflicting allthings-considered oughts, then the possibility of dilemmas has absurd implications. However, as it will soon become clear, there is another characterization of moral dilemmas which, while strong enough to be controversial, does not have such implications. <sup>64</sup> See Connee (1982: 92) and Donagan (1984: 307–8).

conflict seems to have more in common with the structure of desire conflict, since a desire that one decides not to satisfy can survive in a way that a rejected belief cannot. Consequently, Williams concludes, the structure of moral conflict lends support to noncognitivism.<sup>65</sup>

Williams's remarks have since generated a lively philosophical debate. Some have accepted Williams's diagnosis of moral conflict, while denying that it has the metaethical implications he derived from it.<sup>66</sup> Others have argued that conflicts of the sort invoked by him are not genuine moral conflicts. Those who have adopted the former stance have put forward several reasons for doubting that realism rules out the view that conflicting ought statements can simultaneously be true. I have already pointed out that the appearance of inconsistency among such statements can be dispelled by stressing that, whenever two equally stringent moral oughts conflict with one another, the conflict between them is not to be regarded as a conflict between two all-things-considered moral oughts, but merely as a conflict between two non-overridden moral oughts. If what is meant by saying that one has an all-things-considered obligation to adopt a course of action is that the moral reasons in favor of that particular course of action are better than those in favor of any alternative, then it cannot be the case that one has an all-things-considered obligation both to do something and not to do it.<sup>67</sup> Once it is made clear that the conflicting oughts are not all-things-considered moral oughts, there is no reason to see them as inconsistent. As already emphasized, the fact that an act A would have the natural property x may be a reason to see ourselves as morally required to do A, while the fact that an incompatible act B would have the natural property y may be a reason to see ourselves as morally required to do B. Surely, claiming that there are compelling moral reasons to do both A and B involves no contradiction. Note, however, that none of the considerations invoked here has anything to do with the issues that divide realists and anti-realists.

Williams claimed that the appropriateness of moral regret when acting in conflict situations shows that moral conflicts are not soluble without remainder. However, there seems to be another argument to this conclusion which does not turn on whether certain emotional responses are morally appropriate. This argument is present both in Williams's writings and in the writings of those who followed him in claiming that ought statements can genuinely conflict. Let me detail. Williams emphasizes that W. D. Ross's familiar concept of a *prima facie* obligation was called for by situations in which an agent ought to do two incompatible acts. According to Ross, even though two obligations may be present in a situation, only one of them constitutes an actual obligation, while the other one is merely a *prima facie* obligation. Now, let us assume that there are no difficulties in deciding which obligation overrides or defeats the other. What are we to make of Ross's suggestion that overridden obligations are merely *prima facie* obligations? Williams argues that "*prima facie* obligations are not just *seeming* obligations, but more in the nature of the claim, which can generate residual obligations if not fulfilled" (1973*a*: 176, emphasis in original). This seems to have been the idea that Ross had in mind, although

<sup>&</sup>lt;sup>65</sup> See also Williams (1973*b*: 204–5).

<sup>&</sup>lt;sup>66</sup> For arguments to the effect that the existence of genuine moral dilemmas poses no threat to realism, see Guttenplan (1979–80), Foot (1983), Sinnott-Armstrong (1988: 196–200), and Zangwill (1999).

<sup>&</sup>lt;sup>67</sup> See Foot (1983: 385). For a similar point, see also Sinnott-Armstrong (1988: 17–8).

the term "*prima facie* obligation" may have been unsuited to convey it.<sup>68</sup> Elsewhere, Williams adds that:

There are certainly two obligations in a real case of [conflict], though one may outweigh the other. The one that outweighs has greater stringency, but the one that is outweighed also possesses some stringency, and this is expressed in what, by way of compensation, I may have to do for the parties who are disadvantaged by its being outweighed; whether I have merely to explain and apologize, or whether I have to engage further in some more substantial reparatory action. (Williams, 1981: 73–4)

As a number of commentators rightly point out, Ross failed to distinguish between two types of situations. Consider an example that comes from Walter Sinnott-Armstrong (1988: 98–9). If X promises to paint Y's house, but afterwards Y tells X that he no longer needs the house to painted, then X's obligation to paint Y's house is cancelled. By contrast, if X's wife needs to be taken to the hospital at the time when X is supposed to paint Y's house, then X has a justification for not painting Y's house. While in the second case it seems reasonable to assume that X should inform Y about the reason why he did not paint the house and that he should apologize and compensate Y, in the first case X does not owe any apology or compensation. Sinnott-Armstrong emphasizes that it is Ross's failure to distinguish between cases in which an obligation is cancelled from cases in which there is a justification for not fulfilling an obligation that explains why he offers two incompatible descriptions of *prima facie* obligations. On the one hand, by contrasting "prima facie" with "actual", Ross seems to suggest that prima facie obligations are merely apparent. On the other hand, by contrasting "prima facie" with "absolute" he seems to allow that prima facie obligations are real obligations that are not annulled although they are overridden. Yet, given that Ross made use of the concept of prima facie obligations to explain why conflicts of obligations leave a moral residue, it is reasonable to assume that he viewed *prima facie* obligations as real obligations that exert some moral force although they are overridden. In light of these considerations, it may be a good idea to adopt a terminology that departs from the terminology proposed by Ross. It has been common among those who interpret him as claiming that moral conflicts leave a residue to distinguish between prima facie and pro tanto oughts. A prima facie ought is at most an apparent ought, whereas a *pro tanto* ought is a real ought, which may nonetheless be overridden by other oughts.<sup>69</sup>

I have suggested earlier that, even if we do not attribute to proponents of moral dilemmas the view that conflicts between all-things-considered moral oughts are possible, there is no reason to doubt that there is any substantial disagreement between proponents and opponents of dilemmas. We are now in a better position to understand what divides proponents and opponents of dilemmas. Friends of dilemmas share Williams's view that

<sup>&</sup>lt;sup>68</sup> See, for instance, Ross (1930: 20, 28).

<sup>&</sup>lt;sup>69</sup> Those who characterize moral conflict in terms of conflicting moral reasons distinguish between *prima facie* and *pro tanto* reasons. A *prima facie* reason is only an apparent reason, whereas a *pro tanto* reason is a genuine reason, which may nevertheless be overridden by other reasons. *Pro tanto* reasons, as opposed to *prima facie* reasons, are held to have residual reason-giving force. See, for example, Hurley (1985: chs. 7, 9).

moral conflicts leave a moral residue.<sup>70</sup> According to them, it would be a mistake to hold that overridden obligations are annulled. Such obligations exert some moral force, which gives rise to residual obligations. Foes of dilemmas reject this view. For instance, Alan Donagan writes that "the crucial point in Ross's theory is that the prima facie duty with the lesser potential weight has no actual weight at all" (1996: 20). Donagan goes on to insist that "not only do you have no conflict of duties, you do not even have the ghost of a conflict" (ibid.).

So far, I have been considering cases in which one of two conflicting obligations succeeds in overriding the other. Now, let me focus on cases in which one is confronted with two conflicting non-overridden obligations. Given that proponents and opponents of dilemmas disagree about whether the fact that we sometimes seem to be unable to decide whether one obligation overrides another is more than just a symptom of our uncertainty about how to resolve the conflict, let us consider a case that would involve symmetrical choices. Suppose that the lives of two identical twins are threatened and that, although one can save each of them, one cannot save both. Opponents of dilemmas would agree that in this case the moral considerations in favor of taking one course of action do not outweigh the moral considerations in favor of taking the other course of action. According to them, in such cases morality mandates a disjunctive solution. In other words, one has a moral obligation to save either one child or the other. Proponents of dilemmas would also hold that, all things considered, one ought to save one of the children. There is, however, a significant difference between these two positions. On the one hand, foes of dilemmas claim that acting on the all-things-considered ought leaves no moral residue. As Donagan (1984: 307-8) writes, although in a case in which one has to decide which child is to be saved there is clearly a practical conflict, it would be a mistake to think that the conflict in question is a moral conflict. On the other hand, friends of dilemmas hold that in such cases there is a genuine conflict between two moral oughts, and that although all things considered one ought to adopt either one course of action or the other, the original ought that one chooses to disregard retains some moral force. As Nick Zangwill puts it,

The *pro tanto* account lies between two extremes. On one extreme view, we do nothing at all wrong in a dilemma when we act on the overall obligation. [...] This is too weak. On the other extreme view, a dilemma consists of two or more incompatible overall judgments of obligations. This is too strong. There is no reason to believe in such dilemmas. But on the middle way, there are different *pro tanto* obligations which are determined by aspects of the overall situation, and the conjunction of these aspects determines the overall obligation. If we act on the overall obligation we offend at least one *pro tanto* obligation which points the other way. (Zangwill, 1999: 84)

To recap, my suggestion is that, if we look at how different authors deal with the question of whether the resolution of moral conflicts leaves a moral residue, a clear point of contrast emerges between proponents and opponents of dilemmas. This means that we do not need to attribute to proponents of dilemmas the view that conflicts between all-

<sup>&</sup>lt;sup>70</sup> See, for instance, Guttenplan (1979–80), Foot (1983), Marcus (1980: 126, 131–3; 1996: 23), Sinnott-Armstrong (1988: 97–102), Dancy (1993: ch. 7), and Zangwill (1999: 83–4).

things-considered moral oughts are possible in order to conclude that something substantial is at issue in this debate. One caveat is in order here. By claiming that a major point of contention between friends and foes of dilemmas is whether moral conflicts are soluble without a remainder I do not mean to imply that an adequate account of moral dilemmas can be provided solely in terms of the idea of moral residue. It is important to emphasize that, while claiming that the resolution of moral conflicts always leaves a moral residue, one can also hold that it is the possibility of conflicts between nonoverridden moral oughts that raises philosophically interesting questions, and therefore, we should reserve the term "moral dilemmas" for such cases. The perspicuous definition of moral dilemmas is a subject of considerable controversy, and my aim was not to come up with such a definition, but only to show that there is no reason to define moral dilemmas in terms of conflicting all-things-considered moral oughts.

Before concluding this section, let me briefly address a possible objection. One can point out that not all opponents of moral dilemmas share Donagan's view according to which Rossian *prima facie* obligations are not real obligations. For instance, Terrance McConnell - one of the most fervent opponents of moral dilemmas - explicitly allows that overridden obligations exert some moral force (1996: 42). However, while proponents of dilemmas claim that the moral force of an overridden obligation generates an obligation to *apologize* or to *compensate* the parties affected, McConnell argues (a) that this obligation is more plausibly described as an obligation to *explain* one's behavior, and (b) that this "additional obligation" is not generated by the moral force of the overridden *prima facie* obligation, but it is "based on the promotion of fundamental values, such as minimizing unhappiness and showing respect for persons" (1996: 44). In fact, McConnell attempts to preserve Ross's idea that *prima facie* obligations can give rise to residual obligations. So, one can still argue that there is a significant difference between how friends and foes of dilemmas interpret Ross's theory of *prima facie* obligations.

It is also worth emphasizing that it is doubtful whether McConnell succeeds in establishing that overridden obligations do not generate residual obligations. McConnell (1996: 43) asks us to imagine the following example. Suppose that you decide to visit a friend who suffers from depression and that, on your way to his house, you are the witness of an accident and you get involved in rescuing the victim. McConnell argues that, since your friend was not aware of your intention to visit him, you have no obligation to apologize or to explain your behavior. However, this is simply because your friend had no expectations, and not because there is no *prima facie* obligation to visit him. The overridden *prima facie* obligation still exerts some moral force, since it remains true that you ought to visit him at some point. Hence, the suggestion goes, there is no connection between claiming that *prima facie* obligations are not merely apparent and claiming that they give rise to obligations to apologize or to explain still exerts.

Yet, I believe that the reason why we think that no apologies are due in this case is that the obligation to visit a friend who suffers from depression is a good example of what has been called indeterminate or imperfect duties. This means that there are several ways to discharge the duty to help your friend, and that the time and the way in which you discharge this duty are up to you. Thus, the reason why you do not have to apologize or to explain your behavior may also have to do with the fact that you can offer your help on different occasions and you do *not* have a moral obligation to do so on any particular occasion. Certainly, there can be occasions when a friend who suffers from depression needs immediate help and when there is no one else around to help. In such situations, one has a perfect duty to offer one's help. However, McConnell's example is different. Now, suppose that your friend is in such a situation and that you know about it. Surely, if your friend needed immediate help on that particular occasion when you got involved in rescuing an accident victim, it would be less clear that you do not have to apologize or to explain why you did not offer your help even though you knew that it was needed.

Let me take stock. I have argued that *comparative indeterminacy* should not be viewed as inimical to moral realism. If what is puzzling about the idea that moral realism can allow for indeterminacy is the possibility of there being no determinate answer to the question of whether a particular moral property is instantiated in a given case, then comparative indeterminacy poses no threat to moral realism. This is because the abovementioned type of indeterminacy is not due to qualms about whether certain objects of moral assessment would instantiate a given moral property F-ness, but to there being no right answer to the question of how two objects of assessment a and b compare with respect to F-ness.

Still, one might take a different route and claim that comparative indeterminacy is inimical to moral realism because it threatens to lead to inconsistency. Supposing that one is confronted with a choice between two incompatible courses of action, the prospect of comparative indeterminacy leads us to conclude that one may be morally required both to do something and not to do it. At first glance, drawing attention to the fact that moral ties would generate the same problem may seem an easy way out of this difficulty. However, rather than proving that moral realism has no trouble accommodating indeterminacy, this line of thought threatens to turn into an argument against moral realism. Therefore, it seems that the best strategy available to the realist who wants to allow for indeterminacy is to prove that, appearances notwithstanding, claiming that one both ought and ought not to adopt a course of action involves no inconsistency. This is precisely the strategy that I have pursued in this section.

Before going further, one last caveat is in order. Thus far, I have not offered any reason for thinking that there are cases of comparative indeterminacy. My aim was only to show that comparative indeterminacy poses no threat to moral realism. The question whether there is reason to believe that there are cases of comparative moral assessment in which the trichotomy of relations "better than", "worse than" and "equally good" fails will be thoroughly addressed in the second part of the dissertation.

# 2.6.3 Alethic indeterminacy

As we have seen in the previous section, even if we assumed that the possibility of there being no determinate answer to whether a particular moral property is instantiated in a given case is incompatible with moral realism, this would not entitle us to conclude that there are no indeterminate moral issues. This is because in cases of comparative indeterminacy there are no qualms about whether a particular moral property is instantiated. Now, let me turn to the question whether there is sufficient reason to believe that realism cannot allow for the aforementioned possibility. Shafer-Landau (1995) answers this question in the negative. His argument consists of three parts. First, he argues that moral realism does not entail alethic determinacy. As he puts it,

[M]oral realism is committed to the idea that moral judgments are made true by things other than intentional stances taken towards such judgments or towards the facts they convey. Considered in its bare essentials, realism (moral or metaphysical) is neutral with respect to alethic determinacy. *The source of truth conditions and the precision with which they are framed are independent matters*. (Shafer-Landau, 1995: 84, emphasis in original)

Second, Shafer-Landau offers an account of how alethic indeterminacy is generated in cases of noncomparative moral assessment. More concretely, he argues that it is the multidimensionality of moral properties that explains why in some cases there is no determinate answer to the question whether a particular moral property is instantiated. Finally, he emphasizes that opposition to the idea that moral realism is compatible with acknowledging that moral properties can have indeterminate extensions comes from those who insist that vagueness should be viewed as merely an epistemic phenomenon. Yet, Shafer-Landau casts doubt on whether the aforementioned view fares any better than the view that there are vague properties in explaining our irremediable ignorance concerning certain cases that involve property ascription.

Now, let me briefly consider Shafer-Landau's view on how the multidimensionality of moral properties gives rise to alethic moral indeterminacy. Shafer-Landau emphasizes that all moral properties "depend on the satisfaction of a number of distinct constitutive criteria for their instantiation" (1995: 84). According to him, there are several ways in which the multidimensionality of moral properties can generate indeterminacy. First, let us assume that a moral property P depends for its instantiation on the satisfaction of the constitutive criteria A, B and C, and that coming up with an ordinal ranking of A, B and C involves no difficulty. In such cases, if the verdict of the highest-ranked criterion and the verdict of lower-ranked criteria diverge, there may be no determinate answer to whether the moral property in question is instantiated. To illustrate this point, Shafer-Landau considers the example of piety. Admittedly, the criterion that should be accorded the greatest weight in determining the ascription of piety is the sincerity of one's religious belief. However, there are also other criteria that play an important role in determining whether someone is pious. Shafer-Landau asks us to imagine a person who held strong religious beliefs, while at the same time being overcritical, unwilling to spend time doing charitable works, and intolerant of those who do not share his religious beliefs. If one's deeds, and not only one's religious fervor, play a role in determining the ascription of piety, and furthermore, if one's altruism and tolerance are two important dimensions of assessment, then there may be no determinate answer to whether piety can properly be attributed to a person like the one in the above-mentioned example.

Second, Shafer-Landau argues that some borderline cases may result from there being an irremediable puzzle about how to rank the constitutive criteria on which a given moral property depends for its instantiation. Roughly put, the idea is that, if in a certain context the relevant criteria yield conflicting verdicts, and if it is radically unclear which constitutive criterion should be assigned greater weight, then there may be no determinate answer to whether the moral property in question is instantiated.

Finally, Shafer-Landau contends that alethic indeterminacies may be generated by what he calls "criterial vagueness". In the two above-mentioned cases, the assumption was that each of the constitutive criteria that govern the instantiation of a moral property yields a determinate verdict. Yet, Shafer-Landau argues that the constitutive criteria may themselves be multidimensional, and consequently, may generate borderline cases. As he rightly points out, this sort of criterial vagueness can be transmitted to the property whose instantiation depends on the constitutive criteria.

Shafer-Landau (1995: 88) goes on to stress that there is nothing about the structure of alethic moral indeterminacy that would suggest that moral realism is incompatible with this type of indeterminacy. I will consider this claim later on in this section. Before doing that, however, let me first summarize his argument against the view that there can be no vague properties, moral or otherwise.

A predicate is vague if and only if it allows for borderline cases, i.e. cases in which we are in principle unable to know whether a predicate applies or fails to apply. While everyone agrees that there are such cases of irremediable ignorance, the question is what accounts for it. According to a widely held view, our ignorance of whether a predicate such as "bald", "tall" or "red" applies in certain cases is to be accounted for by the lack of precision inherent in our language. Proponents of this view urge that there are no vague properties corresponding to vague predicates. As Shafer-Landau points out, opposition to the idea that there are such properties comes from two camps. On the one hand, there are those who claim that the properties designated by vague predicates should be eliminated from our ontology. According to them, if there really were such properties as baldness or tallness, then they would have perfectly determinate extensions. However, the argument goes, since questions as to whether some person is bald or tall may occasion irremediable puzzlement, we would better give up the idea that there are real properties corresponding to vague predicates.<sup>71</sup> On the other hand, there are those who hold that vague predicates designate real properties that have perfectly determinate extensions. On this view, there is a fact of the matter about whether a predicate like "bald" applies in borderline cases; it is just that we cannot know whether or not it applies.<sup>72</sup> Following Shafer-Landau, I will henceforth refer to the first of the aforementioned views as *radicalism*. The second view is commonly known as *epistemicism* about vagueness.

There is, however, another approach to vagueness. According to some authors, our irremediable puzzlement about whether vague predicates apply to certain cases is best explained by acknowledging that vague predicates have indeterminate extensions. On this view, we are unable to determine whether "bald" or "tall" applies to a certain individuals precisely because there is no fact of the matter about whether certain individuals are bald or tall. As Shafer-Landau puts it, "vagueness is a feature of real properties that is quite properly reflected in the language we use to describe them" (1995: 84). Shafer-Landau stresses that this approach to vagueness provides some advantages over its competitors. As opposed to radicalism, this approach has the advantage of retaining commonsense properties in our ontology. And as opposed to epistemicism, it does not run into the difficulty of having to explain why certain truths about the applicability of vague predicates are epistemically inaccessible to us.

 <sup>&</sup>lt;sup>71</sup> See, for instance, Unger (1979).
<sup>72</sup> See, for instance, Sorensen (1988) and Williamson (1992).

Shafer-Landau points out that a natural response to the view according to which the multidimensionality of moral properties generates alethic moral indeterminacies comes from the proponents of radicalism. Presumably, they would argue that, insofar as multidimensionality gives rise to borderline cases, multidimensional properties have no place in a realistic ontology. As Shafer-Landau emphasizes, this response has certain affinities with a noncognitivist approach to ethics. Thus, one might argue that:

The real world, quite apart from our gilding and staining, contains only discrete physical objects whose nature is free from vagueness. The vagueness enters when we seek to impose moral categories on purely naturalistic phenomena. There are killings but no murders (murders are defined as unjustifiable killings, and "unjustifiable" is anything but precise); promises but no fidelity; acts of giving but no generosity. This is an extension of a familiar response to the sorites, viz., that while there are precise properties of (say) *having 234 hairs on one's head*, or *being 4567 grains of sand*, there really is no such thing as being bald or being a heap. These latter are concepts constructed to satisfy pragmatic demands. (Shafer-Landau, 1995: 91, emphases in original)

Another response comes from those who embrace epistemicism about vagueness. According to them, every statement involving a vague predicate is determinately either true or false. Epistemicists would admit the existence of vague moral predications, but would argue that vague predicates have perfectly determinate extensions.

Yet, Shafer-Landau finds both the above-mentioned views implausible. His main complaint against radicalism can be summarized as follows. If proponents of radicalism are right in claiming that there are no real properties corresponding to vague predicates, then it is difficult to avoid the conclusion that material objects do not exist. Admittedly, any element in the radical's ontology can generate borderline cases. Not only "bald" or "heap" give rise to borderline cases, but also "hair" and "sandgrain". There seems to be no method for removing puzzlement about whether split ends count for more hairs, or about whether freshly shaved hair that can barely been seen with the naked eye counts as hair. One might also identify a unit of measure so small that its gradual diminution would leave us puzzled about whether something counts as grain of sand. Shafer-Landau rightly points out that sorites arguments can be targeted even at subatomic particles. In other words, even in the case of such particles, we can identify a unit of measure so small that its gradual diminution would occasion irremediable puzzlement about whether what we are left with is something distinct from the original particle. Thus, if the generation of borderline cases suffices for ontological eliminability, it follows that there are no material objects. However, this is a conclusion that not many would want to countenance.

Shafer-Landau's complaint against epistemicism is twofold. First, if our irremediable ignorance about the applicability of vague predicates to certain cases is not to be taken as evidence that such predicates have indeterminate extensions, epistemicists must offer an account of *why* we cannot know certain truths concerning the applicability of vague predicates. What's more, since our ignorance in borderline cases is irremediable, what epistemicists have to explain is why we are in principle unable to know such truths. Shafer-Landau points out that, in the absence of a positive argument to the effect that an entire class of truths is beyond our epistemic reach, epistemicism is on shaky ground. He further emphasizes that a metaphysical realist can acknowledge the existence of some

unknowable truths. Yet, he draws attention to the fact that the number of unknowable truths that is required by epistemicism is so large as to cast doubt on the plausibility of this view. Second, Shafer-Landau argues that holding that vague predicates have perfectly determinate extensions leads to absurd consequences. This is because anyone who adopts this view is committed to claiming that statements containing vague predicate are determinately true or false and that there is a single unit that makes the difference between true predications and false ones. On this view, there must be a precise number of hairs that separates bald persons from persons who are not bald. However, Shafer-Landau contends that this view runs into serious trouble. As he emphasizes, whatever the relevant unit, it is infinitely divisible. So, one can object to epistemicism along the following lines:

Suppose that people with fewer than 621 hairs are bald, others not. Suppose I have 622 hairs. Imagine the too-tiring process of slowly removing  $1/10^{20}$ th of a hair. If I continue indefinitely, I will eventually eliminate the hair. If the [epistemicist] is right, there is precisely one point in the process that marks the elimination of the hair. So not only must baldness attributions be sensitive to within one hair; they must be sensitive to within  $1/10^{20}$ th of a hair. (Shafer-Landau, 1995: 93)

In other words, endorsing epistemicism leads to absurd consequences.

In brief, Shafer-Landau points out that a natural response to his account of how alethic indeterminacy is generated in cases of noncomparative moral assessment comes from those who insist that we should resist attributing vagueness to the properties designated by vague predicates. However, he argues that the shortcomings of radicalism, on the one hand, and epistemicism, on the other, provide a "further reason to follow appearances" and to allow that moral realism is compatible with alethic indeterminacy (1995: 93).

Yet, insofar as Shafer-Landau is right in claiming that it is the *multidimensionality* of moral properties that generates borderline cases, it is not entirely clear that what is needed in order to establish that moral realism is compatible with alethic indeterminacy is an argument to the effect that radicalism and epistemicism are untenable. Let me explain.

Shafer-Landau (1995: 84) emphasizes that vagueness typically occurs in terms that are susceptible to sorites arguments. Sorites vagueness presupposes the existence of an underlying scale along which differences between adjacent points are imperceptible. As Shafer-Landau points out, standard sorites properties are *unidimensional*. In other words, there is a single unit of measure whose gradual addition or diminution determines the instantiation of such properties. However, unlike sorites properties, moral properties are *multidimensional*, they depend for their instantiation on a number of constitutive criteria.

At this point, it is useful to distinguish between two types of vagueness, which we may call "sorites vagueness" or "degree vagueness" and "combination of conditions vagueness".<sup>73</sup> The former type of vagueness stems from the lack of a sharp cutoff point along some continuum. By contrast, the latter type of vagueness occurs when there is no determinate answer to the question of what combination of conditions is necessary or sufficient for the application of a term. While sorites vagueness occurs in terms such as "bald", "tall" or "middle-aged", "religion" or "sexual perversion" are examples of terms that exhibit combination of conditions vagueness. For instance, there are a number of

<sup>&</sup>lt;sup>73</sup> For this distinction, see Alston (1964: 87–8).

criteria that are relevant in deciding whether some social practice constitutes a religion. Satisfying all these criteria guarantees that a given practice is a religion, and failing to satisfy any of them guarantees that we are not confronted with a religious practice. Nevertheless, these criteria also leave many cases undecidable. A certain practice may satisfy some but not all of the relevant criteria. If so, we may be irremediably puzzled as to whether the practice in question constitutes a religion. In such cases there is no natural underlying continuum of the sort that characterizes sorites vagueness. Our being unable to tell whether certain practices constitute religions has to do with the difficulties involved in determining which criteria are necessary and sufficient to identify a religion.

Once we make it clear that there are two distinct types of vagueness, it seems quite obvious that moral predicates display combination of conditions vagueness rather than sorites vagueness. For one thing, the *multidimensionality* of the properties designated by moral predicates fits naturally with the former picture of vagueness. Moreover, going back to Shafer-Landau's account of how alethic indeterminacy is generated in cases of noncomparative moral assessment reinforces the idea that moral predicates display combination of conditions vagueness. Think of Shafer-Landau's example involving the ascription of piety. Although one of the criteria that determines the ascription of piety is more salient than the others, borderline cases may arise due to there being no definite answer to whether satisfying this criterion is not only necessary, but also sufficient for the term to apply. In other words, indeterminacy stems from there being no definite answer to the question of what combination of conditions is necessary and sufficient for the application of the term "piety". Cases that involve difficulties in weighing the criteria on which the ascription of a moral property depends raise similar problems. In such cases, our being unable to determine whether or not a moral predicate applies can similarly be explained by our being unable to identify a set of criteria that are necessary and sufficient to determine the application of a given predicate.

Now, the question is whether the fact that moral predicates display combination of conditions vagueness rather than sorites vagueness affects in any way Shafer-Landau's argument. As we have already seen, at least some of the considerations that are supposed to undermine the plausibility of the view that there can be no vague properties are specifically targeted against holding that statements containing sorites predicates must be determinately true or false. More specifically, Shafer-Landau argues that epistemicism implausibly suggests that every case of puzzling predication can be resolved and that it is an infinitesimal unit that makes the difference between true predications and false one. It important to note, however, that someone who maintained that, for any is (multidimensional) predicate P, there must be a set of conditions that are necessary and sufficient for the predicate to apply (or that holding otherwise would be inimical to realism) would not be committed to absurd claims of the sort mentioned by Shafer-Landau. That is, someone who adopted the above-mentioned view would not be committed to the claim that it is the presence or absence of an infinitesimal, and thus imperceptible, unit that makes the difference between true predications and false ones. Telling apart cases in which all the conditions that are relevant in determining the application of a given predicate are satisfied from cases in which only some of these conditions are satisfied would involve no serious difficulty.

Yet, one can argue that in order to establish that moral realism is compatible with alethic indeterminacy, all that one needs to do is to dispel the idea that realists about any

domain are committed to thoroughgoing determinacy. Furthermore, one can point out that our answer to the question whether realism entails such a commitment to determinacy partly turns on which of the above-mentioned approaches to vagueness is more tenable. As already emphasized, there are two broad alternatives to the view that vague predicates have indeterminate extensions, and the rejection of at least one of them is driven by considerations peculiar to vagueness of the sorites variety. Thus, it seems that dealing with sorites vagueness may help settle the debate over whether moral realism can allow for alethic indeterminacy even if this is not the type of vagueness that moral predicates typically display.

To recap, insofar as resistance to the idea that moral realism can accommodate alethic indeterminacy is just an expression of the view that reality is fully determinate, and consequently, any question concerning the instantiation of a real property must have a determinate answer, Shafer-Landau's strategy seems adequate. Still, some might insist that there are two distinct types of vagueness, each raising problems interesting in their own right. So, given that moral predicates exhibit combination of conditions vagueness rather than sorites vagueness, those who argue for the compatibility between moral realism and alethic indeterminacy should address the question whether the former type of vagueness poses a problem for realism (i.e. a problem that attaches specifically to this type of vagueness).

What reason is there to think that realism cannot accommodate indeterminacy with respect to what combination of conditions is necessary or sufficient for certain terms to apply? As already pointed out, while the application of sorites predicates is governed by the natural ordering of cases on an underlying continuous scale, in the case of predicates that display combination of conditions vagueness there is no such natural continuum. The application of such predicates depends on the way we group things together according to certain similarities. A crucial question arises at this point, namely, whether reality naturally breaks up into our classifications. Suppose this question is to be answered in the negative. Then, indeterminacy in the application of our terms can readily be explained by reference to the fact that the application conditions for these terms depend on human intentionality. More specifically, it is the incompleteness of our intentions that explains indeterminacy. In this case, indeterminacy is plainly incompatible with realism, since to admit that the application conditions for predicates in a given class are determined by our intentional attitudes means to admit that the properties designated by the predicates in question are not to be construed realistically. What if our systems of classification mirror what is "out there"? Some will undoubtedly insist that, insofar as our classifications can be viewed as answerable to the outside world, failures of determinacy with respect to what combination of conditions is sufficient for certain terms to apply are attributable to our epistemic limitations. We have seen above that claiming that any case of puzzling predication due to combination of conditions vagueness can be resolved does not lead to the same counterintuitive consequences as claiming that there is always a determinate answer to questions concerning the applicability of sorites predicates. Yet, the question whether realism mandates the former view needs to be discussed at greater length. I will return to this question later on. The important thing to note here is that focusing on combination of conditions vagueness brings out more clearly the way in which realism is threatened by vagueness. Combination of conditions vagueness occurs in terms that have a number of logically independent conditions of application. Given that what is required to determine the applicability of such terms is tracing similarities and differences between things, the question whether our classifications correspond to "real properties" or merely reflect our own interests is raised in a particularly acute way.

Now, the question is whether moral realists have the resources to establish that some moral predicates have indeterminate extensions. In other words, what we need to know is whether there is sufficient reason to conclude that the aforementioned view is to be preferred both to the view that there are no real properties corresponding to vague moral predicates and to the view that there are real moral properties, but they have perfectly determinate extensions. Consider, first, the view that moral properties have no place in a realistic ontology. If the existence of borderline cases due to combination of conditions vagueness were sufficient to create a case against realism about a given domain, then obviously anti-realism would not be restricted to the moral domain, since moral predicates are not the only ones that exhibit this type of vagueness. However, this does not mean that the aforementioned type of vagueness threatens to lead to global antirealism. One can, for instance, point out that having several distinguishable conditions of application is in itself insufficient to render a predicate vague. As William Alston (1967: 219) emphasizes, even though the predicate "square" has two conditions of application (i.e. being a rectangle and having all sides equal), there is a determinate answer to the question of what combination of conditions is necessary and sufficient for the predicate to apply. The answer is that each of the above-mentioned conditions is necessary but not sufficient for its application, and that their combination is sufficient for "square" to apply. By contrast, there seems to be no determinate answer to the question of what combination of conditions is necessary or sufficient for the application of a predicate like "religion". Since it appears that not all multidimensional predicates are affected by combination of conditions vagueness, one might hope to prove that those domains about which most of us want to remain realist are free of vagueness. Thus, one might contend that social and institutional terms are the most susceptible to combination of conditions vagueness, while at the other extreme we find natural kind terms such as "gold", "water" or "whale", which are supposed to be immune to vagueness. Furthermore, one might add that this difference should properly be reflected in our taking a realist stance towards some domains, but not towards others.

Admittedly, indeterminacy in the application of certain social or institutional terms can be traced back to the fact that the applicability of such terms depends on what conditions we collectively accept as sufficient for their application. This is perhaps most obvious in the case of institutional terms. Consider an example that comes from Amie Thomasson:

[S]uppose that according to election rules, if a ballot in a presidential election in Florida contains a hole punched by a candidate's name, that is a vote, and if no hole is punched, then that is not a vote. In such cases, it clearly becomes a matter of stipulation, not discovery, whether or not to "count" hanging chads and their like as cases of votes. Since these are explicitly institutional terms, it should be no surprise that their application conditions are determined by the conditions we accept as relevant to their application, so if there are cases that (e.g.) the stipulators have not thought of, and not considered whether or not the term applies, there is no fact of the matter about whether or not the term applies in those conditions. (Thomasson, 2007: 93)

In other words, to the extent that the stipulators have not considered certain questions regarding the applicability of the term "vote" (such as whether hanging or dimpled chads count as votes, or whether double-punched ballots, or ballots containing both a clear punch and a dimple in the chad of a different presidential candidate count as votes), there is no determinate answer to these questions. Natural kind terms stand in sharp contrast to institutional terms, for the application conditions of such terms do not seem to depend on our beliefs or practices. This point can be expressed by saying that scientific conclusions about the conditions that are necessary and sufficient for the application of a natural kind term are a matter of discovery, not stipulation. Now, if we stick to what many take to be the paradigmatic example of natural kinds, i.e. chemical kinds, a strong case can be made that terms which designate chemical elements and chemical compounds are not affected by combination of conditions vagueness. This might suggest that there is a correlation between a term's giving rise to vagueness and the fact that its application conditions are dependent on us (or between a term's being immune to vagueness and the fact that its application conditions are independent of us). Put differently, the idea is that the fact that the terms associated with a given domain exhibit combination of conditions vagueness should prompt us to take an anti-realist view of the domain under consideration. Thus, given that moral terms display the aforementioned type of vagueness, we apparently have reason to abandon moral realism.

This is, however, too quick. For one thing, not all natural kind terms are immune to combination of conditions vagueness. Consider biological species, which were traditionally thought to be paradigmatic natural kinds. As it is commonly emphasized in the literature on natural kinds, species terms allow for borderline cases.<sup>74</sup> Indeterminacy in the application of species terms is explained by reference to the fact that species cannot be individuated on the basis of a set of necessary and sufficient characteristics. Yet, great care needs to be exercised in spelling out the aforementioned claim. Empirical research shows a high level of intraspecific morphological and genetic variability.<sup>75</sup> In fact, neither morphological nor genetic features are likely candidates for being the essential properties of biological species, since there are no morphological or genetic features that are shared by all members of a species. Nevertheless, we are not entitled to conclude that biological species have no essential properties. The important thing to note here is that, actually, biologists do not delimit species on the basis of intrinsic properties. The criteria used to individuate species are relational or historical. For instance, according to some biologists, species are to be viewed as groups of interbreeding natural populations which are reproductively isolated from other such groups. According to others, species are individuated on the basis of shared ancestry. Thus, one can argue that the view according to which biological species have essential properties is not threatened by the fact that there are no intrinsic properties that are shared by all members of a species. More to the point, one can claim that relational or historical properties can serve as the essential properties of biological species.<sup>76</sup>

Yet, even if we focus on relational or historical properties, it is far from clear that there is a set of properties that are necessary and sufficient for species membership. As it

<sup>&</sup>lt;sup>74</sup> See, for instance, Boyd (1999; 2006). <sup>75</sup> See, for instance, Dupré (1981: 84–5).

<sup>&</sup>lt;sup>76</sup> See, for instance, LaPorte (2004: 64–5).

will soon become apparent, using relational or historical criteria for species membership does not ensure that species have clear boundaries. To be sure, the two above-mentioned approaches delimit species differently. In other words, some organisms may belong to a species according to one approach, but not according to the other. However, this does not mean that biological species have no clear boundaries. For one can argue that, although biologists propose different ways of classifying organisms, at most one of these systems of classification corresponds to real divisions in nature.<sup>77</sup> Therefore, the argument would go, the most pertinent question is whether the boundaries of biological species are indeterminate according to that system of classification which mirrors ontological reality. Now, let us suppose that classification by species does reflect real divisions in nature and that species consist of populations of organisms that are able to successfully interbreed. Then, difficulties would arise in determining whether certain morphologically similar populations of organisms are potentially capable of interbreeding. Alternatively, suppose that species are to be individuated on the basis of shared ancestry. According to this approach, species evolve gradually and, at some point during the process, some members may diverge from the rest of the population. When the divergence between two such groups becomes sufficiently clear, they can be viewed as separate species. It should be emphasized, however, that actually determining when a new species comes into being is more complicated than it may at first seem. As Richard Boyd (1999) rightly points out, indeterminacy in the application of species terms seems an unavoidable consequence of evolutionary theory. Boyd follows Darwin's lead in claiming that speciation presupposes the existence of populations that are intermediate between the mother species and the newly arisen one. He writes that "any 'refinement' of classification which artificially eliminated the resulting indeterminacy in classification would obscure the central fact about heritable variations in phenotype upon which biological evolution depends, and would thus undermine the accommodation of the classificatory resources of biology to relevant causal structures" (1999: 68).

Where does this leave us? We have seen that not all natural kind terms are immune to vagueness. As emphasized above, species terms allow for borderline cases. Thus, one must either give up realism about species and hold that the boundaries of biological species are fixed by us rather than by nature, or else accept that realism about species is not wedded to the view that there is a set of properties whose possession is necessary and sufficient for species membership. Being a realist about biological kinds would involve construing them as dynamic rather than static, thus contradicting the orthodoxy about natural kinds.<sup>78</sup> Yet, one can argue that the lesson to be drawn from the case of biological species is not that the orthodoxy should be rejected. Perhaps what is important is that some of the kinds studied by our natural sciences seem to conform to this orthodoxy. Chemical kinds, for instance, do not exhibit the dynamism of biological species. Also, it is quite uncontroversial that having a particular atomic number is necessary and sufficient for an atom to be a sample of a certain chemical element (e.g. having the atomic number 79 is necessary and sufficient for an atom to be a sample of a semple of gold). So, why not conclude that, while chemical kinds are genuine natural kinds, biological species are not?

<sup>&</sup>lt;sup>77</sup> For a different view, see Dupré (1981). Dupré defends what he calls "promiscuous realism", according to which there are different equally defensible ways of classifying organisms.

<sup>&</sup>lt;sup>78</sup> For more on this point, see Boyd (1999).

Two points are worth emphasizing here. First, if we insist that realism about natural kinds involves viewing them as static and eternal and maintaining that there are necessary and sufficient conditions for kind membership, then many of the kinds that constitute the object of special sciences would fail to count as genuine natural kinds. To see that this is so, one need only consider the case of geological sciences or meteorology. Second, there is a powerful argument against holding that only groupings that conform to the foregoing picture count as genuine natural kinds. As Boyd (2006) emphasizes, natural kinds are identified as those kinds that support inductive and explanatory inferences.<sup>79</sup> Roughly put, the idea is that what accounts for the success of our inductive and explanatory practices is precisely the existence of natural kinds. According to Boyd, natural kinds are homeostatic clusterings of properties. That is to say, they are the result of a causal homeostasis, where the presence of some properties tends to favor the presence of others, or where there are underlying mechanisms that tend to maintain the presence of the clusterings. Thus, one can say that, unless there were such causally stable clusterings of properties, no inference from the presence of some such properties to the presence of other properties could be epistemically reliable. Some of the consequences of viewing natural kinds as homeostatic clustering of properties will be analyzed later on. What interests me here is simply the view that the naturalness of natural kinds consists in their being suited for induction and explanation. Boyd points out that "the phenomenon that the theory of natural kinds explains – successful inductive and explanatory inferences. and the accommodation of conceptual resources to the causal structures that underwrite them - occurs no less in inductive/explanatory enterprises that seek (and achieve) more local and approximate knowledge than in fundamental physics" (2006: 412). The fact that we are able to identify causally sustained regularities in disciplines such as biology, geology or meteorology should prevent us from concluding that there are no natural kinds in such disciplines. As Boyd rightly stresses, even if the regularities identified in these disciplines are not eternal or exceptionless, natural kinds are still needed to explain what it is about the world that makes inferential practice in these disciplines epistemically successful<sup>80</sup>

Let me take stock. My starting point was the observation that indeterminacy in the application of certain social or institutional terms can readily be explained by reference to the fact that the applicability of such terms depends on what conditions we collectively accept as sufficient for their application. I suggested that those who believe that realism is committed to thoroughgoing determinacy might contrast social and institutional terms with natural kind terms, whose application conditions do not seem to depend on us. The point of the contrast would be to establish that whether a term gives rise to vagueness is simply a function of whether its applicability is dependent on human intentionality. Yet, I have argued that not all natural kind terms are immune to vagueness. As we have seen,

<sup>&</sup>lt;sup>79</sup> Quine (1969) was the first philosopher to suggest that what makes our inductive inferences epistemically reliable is the existence of natural kinds.

<sup>&</sup>lt;sup>80</sup> Some might object to characterizing natural kinds in terms of there being suited for induction and explanation on the grounds that things that intuitively fail to constitute a kind may participate in causal regularities. For instance, even though the fact that positively charged things repel one another seems to constitute a law of nature, it can be argued that positively charged objects do not form a natural kind. To this Boyd would reply that "the naturalness of a natural kind – its suitability for induction and explanation – is *discipline relative* "(2006: 411, emphasis in original). As he writes elsewhere, "we can and must 'cut nature at its joints', but the boundaries between joints are themselves context-specific" (1993: 514).

species terms allow for borderline cases. Moreover, we have seen that (a) indeterminacy in the application of species terms is ineliminable, and that (b) the naturalness (and hence the "reality") of biological kinds is not threatened by this indeterminacy. The upshot is that the existence of borderline cases due to combination of conditions vagueness is insufficient to create a case against realism about a given domain. Thus, the mere fact that moral terms display the aforementioned type of vagueness does not justify taking an antirealist stance towards morality.

The same conclusion can be reached by a slightly different route. Those who hold that realism cannot accommodate indeterminacy can stress that the existence of kinds that have natural boundaries presupposes that certain ignorance and error principles apply. The idea that a kind has natural boundaries can be expressed by saving that it has its "boundaries guite independently of how our concepts and representations might happen to divide things up, in particular, independently of what we believe about the conditions relevant to drawing those boundaries" (Thomasson, 2003: 582). To put the matter slightly differently, the conditions that are relevant to determining whether something is a member of a natural kind are independent of whether anyone accepts those conditions. As Thomasson (2003: 583) emphasizes, what follows is that it is possible that the conditions that determine kind membership remain unknown to everyone. Thomasson calls this the "Ignorance Principle". Furthermore, she points out that, since the boundaries of natural kinds are not determined by our beliefs about them, then any beliefs about how to draw these boundaries can turn out to be massively wrong. To illustrate this second principle, which Thomasson calls the "Error Principle", consider the once-popular belief that whales are fish. Thomasson (2003: 590, n12) rightly stresses that, while such massive errors about the nature of a kind are possible in the case of natural kinds, the same does not hold in the case of social or institutional kinds. The two above-mentioned principles might be used to argue that, even if certain questions about whether some entities belong to a natural kind leave us irremediably puzzled, this does not mean that there is no determinate answer to such questions. Those reluctant to accept that realism can allow for indeterminacy would most probably insist that, whenever a question concerning kind membership seems to admit of no determinate answer, this appearance can be explained away by reference to the aforementioned principles.

It is important to note, however, that those who claim that indeterminacy in the application of natural kind terms reflects an indeterminacy in the natural world argue that refinements of classification would only render the definitions of natural kinds unnatural. As Boyd puts it, such refinements "would either require that we treat as important distinctions which are irrelevant to causal explanation or to induction or that we ignore similarities which are important in just these ways" (1988: 198). Thus, on this view, our being puzzled about whether a natural kind term applies to a certain case is simply due to there being no way of remedying the indeterminacy while, at the same time, remaining truthful to the aim of employing categories that correspond to the causal structures of world. This view offers a plausible answer to a common objection against holding that the principle of bivalence should be abandoned. According to this objection, the claim that a sentence is neither true nor false needs as much positive argument as the claim that a sentence is either true or false. However, the objection goes, this is precisely what those who claim that the principle of bivalence should be abandoned fail to provide.<sup>81</sup> Now, if

<sup>&</sup>lt;sup>81</sup> See, for instance, Dworkin (1996*a*).

Boyd is right in holding that remedying the indeterminacy in the application of natural kind terms would involve misrepresenting the causal structures of the world, then it is reasonable to conclude that certain sentences containing natural kind terms are neither true nor false. This conclusion is not based on there being no evidence pointing either way (i.e. on our *not knowing* whether certain sentences are true or false), but on judging that to maintain that certain sentences are true or false would be scientifically misleading.

Surely, those who contend that realism is incompatible with indeterminacy might argue that, in judging that indeterminacy in the application of natural kind terms cannot be remedied without misrepresenting the causal structures of the world, we are in fact committing errors of the kind predicted by the "Error Principle". Also, they might raise doubts about whether biological species, which are used by Boyd to illustrate the point that indeterminacy is ineliminable, count as genuine natural kinds. Yet, such worries can be answered by appealing to considerations about the success of our inductive practices. To the extent that our beliefs about the nature of a kind are the result of well-confirmed inductions, there is little reason to suppose that such beliefs will turn out to be massively wrong. And such well-confirmed beliefs often support the conclusion that the question whether to treat some entities as members of a given kind has no determinate answer. The worry that biological kinds do not count as genuine natural kinds has been addressed above. As we have already seen, the success of our inductive practices in disciplines such as biology, geology or meteorology should prevent us from concluding that there are no natural kinds in such disciplines. What seems to distinguish natural kinds from the kinds that are the result of "gerrymandering" is the fact that only the former are well-suited to support inductive inferences. As Kathrin Koslicki rightly points out, "what is deficient about [gerrymandered classifications] is precisely [...] that their members lack any other common characteristics (or at least ones that are not themselves gerrymandered) besides the features by means of which they are categorized under a common heading" (2008: 791).

This concludes my argument against holding that the existence of borderline cases due to combination of conditions vagueness is sufficient to create a case against realism about a given domain. The upshot is that the mere fact that moral terms display the above-mentioned type of vagueness does not justify moral anti-realism.

In the Appendix, the issue of whether indeterminacy is compatible with realism will be approached from a different angle. Let me detail. At the beginning of this section, I have considered Shafer-Landau's account of how alethic indeterminacy is generated in cases of noncomparative moral assessment. According to him, it is the multidimensionality of moral properties that is at the root of alethic indeterminacy. Shafer-Landau argues that the instantiation of moral properties depends on the satisfaction of several constitutive criteria and that puzzlement with regard to whether a given moral property is instantiated may arise due to the difficulties involved in weighing the conflicting verdicts yielded by determinately ranked constitutive criteria or to the difficulties involved in weighing the criteria themselves. However, those who want to resist the idea that moral realism allows for alethic indeterminacy can respond that none of this implies that moral properties have indeterminate extensions. More specifically, they can point out that the weighing of the constitutive criteria (or the weighing of the conflicting verdicts yielded by determinately ranked constitutive criteria) is supposed to be determined by our concept of the moral property in question, and that even when this is not the case there is no *a priori* reason to suppose that the question whether a particular moral property is instantiated admits of no right answer. Perhaps an example will help to drive this point home.

Consider an example that comes from János Kis (2002). According to a widely held view, an act of killing counts as killing in self-defense if and only if two conditions are met. Let M represent the person who uses deadly force and N represent the person against whom deadly force is used. M can be said to act in self-defense if and only if (a) N poses an *immediate* threat to cause death or serious bodily harm to M, and (b) the use of deadly force is *necessary* to avert the threat posed by N. Now, the question is whether a case in which a battered woman kills her husband ought to be treated as a case of killing in selfdefense. Admittedly, this is not a standard case of self-defense. A battered woman will typically kill her abusive husband while he is drunk or asleep. The fact that the battered woman acts in a preemptive fashion counts against treating this case as one of killing in self-defense. However, as Kis rightly emphasizes, the point of the immediacy condition is to ensure that the victim has no other options than to defend herself by using force. The idea is that a person who is threatened with grave and imminent harm is exempted from the rule that only the state's coercive agencies can legitimately use violence. It should be kept in mind, however, that a woman who is systemically isolated from the outside world by her husband may be deprived of access to the state's protective agencies. Thus, even if the immediacy condition fails to obtain, there are considerations that count in favor of treating the killing of the batterer as a case of killing in self-defense.

What is undeniable is that there is no easy answer to the question of whether a case like the one discussed above falls within the extension of the concept of self-defense. And one can argue that our difficulties in determining whether or not this is the case have to do with the fact that, when we devised the concept of self-defense, we had in mind only a certain kind of situations, and we framed the criteria for the applicability of the concept with respect to such situations. No wonder, then, that these criteria fail to unequivocally determine whether the battered woman case ought to be treated as one of killing in selfdefense. Much the same seems to hold for most of our evaluative concepts, i.e. when it comes to unenvisaged situations, a concept's applicability criteria may yield no definite answer to whether it can properly be applied. However, even if our previous definition of a concept fails to yield a definite answer to whether the concept in question can properly be applied to a novel situation, this does not mean that such an answer is in principle unavailable.

One way to substantiate this point would be the following. Arguably, the idea that in unforeseen circumstances we may be genuinely puzzled about whether or not to apply a hitherto well-behaved concept has its origins in the writings of Friedrich Waismann. In his *Verifiability* (1945), Waismann pointed out that most of our empirical concepts are *open textured*, i.e. even if they are not currently identifiable as vague, in unanticipated circumstances their application may become indeterminate. He called attention to the fact that, when we define an empirical term, we frame its application conditions with respect to foreseeable circumstances. However, as Waismann rightly emphasized, the possibility that we have not taken into account something that is relevant to the term's usage can never be eliminated. It is along these dimensions which have not been considered that the application of our terms might prove indeterminate. Moreover, when several conditions are relevant to determining the applicability of a term, they may yield divergent verdicts, thus leaving us puzzled about whether the term can legitimately be applied. One of the questions that deserve a more detailed treatment than I am able to provide here is to what extent we are warranted in extending Waismann's conclusions about empirical concepts to moral concepts. I will nevertheless assume that applying Waismann's analysis to moral concepts poses no special difficulty and focus instead on two other questions: (a) whether indeterminacy of the open texture variety is supposed to be ineliminable, and (b) whether holding that open texture is a fundamental characteristic of our concepts, I believe that a careful examination of Waismann's view concerning the open texture of empirical concepts will throw considerable light on the two above-mentioned questions.

Once we focus on Waismann's view according to which most empirical concepts are open textured, those who want to resist the idea that realism allows for indeterminacy can make two points. First, they can argue that Waismann should not be read as claiming that indeterminacy of the open texture variety is ineliminable. This argument might go as follows. To claim that a term is open textured is to claim that the possibility that we have not taken into account something that is relevant to the term's usage, something that would eventually make us extend or modify its previous definition, is always present. It is important to note, however, that the emphasis is not on the fact that even our best current theories might prove to be wrong or incomplete. Instead, the emphasis is on the fact that we cannot possibly envisage all the circumstances in which our empirical terms might be put into test, and that even if we can conceive of a number of such circumstances, it would be unreasonable to build decisions about all such cases into the meaning of our terms. The reason why it is unreasonable to settle in advance certain questions regarding the applicability of our terms is that doing so would require making decisions about what beliefs we would adopt if our current theories proved wrong, and making such decisions on the basis of our linguistic intuitions does not seem justified. To put it more succinctly, the idea that Waismann was trying to convey by the metaphor of open texture was not that actual indeterminacy cannot be remedied, but rather that potential indeterminacy is ineliminable.

Second, those who maintain that realism and indeterminacy are incompatible might argue that the thesis according to which in unanticipated circumstances the application of most of our empirical terms may become indeterminate is licensed by a particular view of language, one that metaphysical realists rightfully reject. More concretely, they can point out that, on a picture of language that equates meaning with use, the application of our terms becomes indeterminate in unusual cases precisely because their previous use offers no clear guidance about the correct way to go on. However, the argument would go, on the picture of language advanced by metaphysical realists, even if we were uncertain about whether a given term applies in a particular case, it would be misleading to say that the application of the term is indeterminate. According to the so-called causal theory of reference, determining the meaning of a term is amenable to empirical investigation. The proponents of the causal theory stress that the correct definition of a term like "gold" or "whale" is an *a posteriori* matter (or, to put it differently, the question of how to define "gold" or "whale" cannot be settled by means of conceptual analysis). Yet, insofar as determining the meaning of such terms involves investigating the features of the world,

and insofar as their definitions are essentially revisable, the mere fact that nothing in the *previous* use of a term dictates whether the term can properly be applied to an unfamiliar case is insufficient to establish that its application is indeterminate.

In the Appendix, I will offer a brief response to the two lines of argument presented above. I will attempt to prove that metaphysical realists can accommodate the claim that open texture is a pervasive feature of language even if this claim is taken to imply that *actual* indeterminacy cannot be remedied.

# 2.6.4 Response-dependence and indeterminacy

Before moving on to the next chapter, one last point of clarification. Thus far, I have dealt with the question of whether moral realism, characterized in terms of mind-independence, is compatible with indeterminacy. Recall, however, that at the beginning of this chapter I have put forward another characterization of moral realism. According to it, any view which countenances the existence of robust moral facts and properties counts as properly realist. On this characterization, response-dependent accounts of moral facts and properties qualify as realist.<sup>82</sup> Yet, it is important to emphasize that, if moral realism is characterized in terms of mind-independence, then holding that moral facts and properties are entirely independent of characteristic human responses is necessary for a view to qualify as realist. So the two above-mentioned characterizations of moral realism are not extensionally equivalent. In what follows, I will make a few brief remarks about the relationship between response-dependence and indeterminacy.

Advocates of the response-dependent view hold that moral properties are analogous to secondary qualities (colors, in particular). On this view, to claim that something has a certain moral property is to claim that it is such as to elicit a certain type of response from normal subjects under normal circumstances. As we have already seen, the question of whether response-independent versions of moral realism can allow for indeterminacy has been the object of a long-standing philosophical controversy. It is important to notice, however, that there is a fair amount of agreement about the fact that response-dependent theories would introduce new sources of indeterminacy into the picture. Those who argue for a response-dependent account of morality regard indeterminacy as an unavoidable corollary of their view. For instance, Pettit writes that:

Suppose that there is a substance such that when it is exposed to a photon of light it changes in a manner that affects how it appears, even appears to normal subjects. Before exposure, as we might incautiously say, it was disposed to look green; after exposure, it is disposed to look red. Is the object really green or really red? Someone who adopts a response-dependent line will naturally take the view that this is a borderline case [...]. (Pettit, 1991: 618)

To claim that the property of being red is response-dependent is to claim that an object is red just in case it appears red to standard perceivers under standard circumstances. Now, the question whether the object in the above-mentioned example is really red admits of

<sup>&</sup>lt;sup>82</sup> See *supra* notes 4–5.

no determinate answer precisely because, as Pettit puts it, "the response that is associated with color is not forthcoming in the ordinary way" (ibid.). Quite to the contrary, in such a case "the regular practice of determining color is systematically thwarted" (ibid.). And the fact that our responses are systematically thwarted is all that is needed to establish that there is no determinate answer to whether the object in question is green or red. On the response-dependent view, there can be no question of looking deeper into the matter.

The question that arises at this point is whether construing the facts in a given domain as response-dependent implies that whatever indeterminacy there may be in the application of the concepts associated with the domain in question should be attributed to these concepts' applicability being dependent on us. Our answer to this question will turn on how we answer a different question that is often asked in the literature on responsedependence.<sup>83</sup> When a property P is associated with characteristic human responses, one might ask a Euthyphro-like question. The question raised by Socrates in Plato's dialogue *Euthyphro* (9D–11B) is whether certain acts are pious because they are loved by the gods, or whether the gods love such acts because they are pious. In the same vein, one might ask whether something is P because it elicits a certain response from standard subjects, or whether it elicits a certain response from standard subjects because it is P. Advocates of response-dependent theories typically claim that the "because" involved here can run in both directions.<sup>84</sup> On the one hand, we can say that something is red because it appears red to standard subjects. But on the other hand, we can say that it appears red to standard subjects because it is red. As Pettit puts it, the property of being red is what ensures that under suitable circumstances standard subjects will have a certain visual experience.

Assuming that advocates of response-dependent theories are right to point out that adopting a response-dependent account of colors does not imply that color properties can retain no objective footing, let us consider how this bears on the question of whether or not indeterminacy in the application of our color concepts mirrors an indeterminacy in the external world. Response-dependent theorists argue that color properties pull their weight in the explanation of color responses. According to them, an object's being disposed to look red is what causes red experiences in normal subjects under normal circumstances. It is important to note that, on this view, the extension of color properties is not fixed by us. To put the point more concretely, we can think of ourselves as detecting the extension of color properties, although the extension of these properties is determined in such a way that only beings who share our perceptual capacities can identify it. If we are tracking or detecting the extension of color properties, and if agreement in color responses is to be taken as evidence of this fact, then it seems reasonable to suppose that in cases like the one mentioned by Pettit, when our regular practice of determining color is systematically thwarted, reality is silent as to whether a certain color concept can properly be applied.

Much the same would hold for moral properties. Those who advocate a responsedependent account of morality claim that, if P is a moral property, an object's being P is not exhausted by our judging it to be P. According to them, an object's being P is what explains the fact that we judge it to be P. In other words, an object's having a certain moral property ensures that under suitable circumstances it will elicit certain judgments from suitable subjects. Thus, response-dependent theorists can point out that, since moral judgments track moral properties, and since agreement in moral judgment can be taken as

<sup>&</sup>lt;sup>83</sup> See, for instance, Johnston (1989: 171), Pettit (1990: 18; 1991: 614), and Wright (1992: 108–139).

<sup>&</sup>lt;sup>84</sup> See, for instance, Wiggins (1987*a*: 106–7) and Pettit (1991: 614–5).

evidence of this fact, it seems reasonable to suppose that, when there is no prospect of securing agreement among suitable subjects, despite the conditions of investigation being suitable, the best explanation of this fact might be that there is no fact of the matter.<sup>85</sup>

I have not offered here the sort of detailed argument that would be required in order to conclusively establish that response-dependent accounts of morality succeed in making plausible the idea that indeterminacy in the application of our moral concepts reflects an indeterminacy in the external world. Undoubtedly, those who adopt this position need to go to greater length in explaining why the attempt to elucidate the relationship between property and response in terms of a two-way dependence relationship involves no vicious circularity. Furthermore, they need to offer a convincing argument to the effect that moral properties are response-dependent in the requisite sense. This might involve, for instance, putting to rest certain worries about whether the analogy between secondary qualities and moral properties holds up.<sup>86</sup> My aim, however, was merely to draw attention to the fact that, if it turns out that moral properties are to be conceived of as response-dependent, it does not automatically follow that whatever indeterminacy there is in the application of moral concepts is attributable to these concepts' applicability being (in some sense to be specified further) dependent on us.

In this chapter, I have argued that response-independent versions of realism have no trouble accommodating indeterminacy. So one might be tempted to think that, if it turns out that those versions of realism which purport to deliver a stronger form of objectivity have the resources to accommodate indeterminacy, then there is even less reason to be concerned that response-dependent versions of realism, which deliver a weaker form of objectivity, are incompatible with indeterminacy. Still, a worry remains as to whether response-dependent theories can even make sense of the idea that indeterminacy in the application of moral concepts mirrors an indeterminacy in the external world. The above remarks were intended to dispel this worry.

Thus far, I have not provided any reason for thinking that response-dependent theories should count as properly realist.<sup>87</sup> Insofar as the proponents of such theories are able to rebut the charge of circularity, one might argue that there is a legitimate sense in which response-dependent properties are real properties of objects (since such properties would be needed to explain our responses). Yet, the question of what exactly is required for a metaethical view to qualify as realist is a controversial question that I will not deal with at this point. As has been frequently pointed out, the debate over whether response-dependent accounts of morality should count as realist tends to turn on what is required for realism in general rather on issues having to do with the nature of morality.<sup>88</sup> So, offering an adequate treatment of this question would take me too far afield. Note, also, that it is not essential to my argument that response-dependent theories be considered realist. The proponents of the argument from disagreement maintain that the distinctive nature of moral disagreement cannot be explained consistently with the view that there

<sup>&</sup>lt;sup>85</sup> See Wiggins (1987*b*: 149; 1990–91).

<sup>&</sup>lt;sup>86</sup> For an argument to the effect that the analogy between secondary qualities and moral properties is flawed, see Wright (1988). For highlighting certain respects in which color is disanalogous with moral properties, see also Blackburn (1985*a*).

<sup>&</sup>lt;sup>87</sup> The question whether response-dependent accounts of moral properties should count as realist has been the object of philosophical controversy. See, for instance, Dancy (1986), Pettit (1991), Smith (1993), and Wright (1993).

<sup>&</sup>lt;sup>88</sup> See, for instance, Hookway (1986: 202).

are moral facts. As I have already emphasized, if an explanation of disagreement consistent with a weaker form of moral objectivism were available, the argument from moral disagreement would be refuted. In answering to the challenge posed by moral disagreement, the important thing is that the objectivist proves that his argument does not trade on intuitions that support different versions of moral objectivism.

Some might argue that there is an important respect in which moral properties are disanalogous to colors, i.e. while it is usually assumed that color properties figure in the causal explanation of our color experiences, a causal model of moral properties is highly problematic. Moral properties - or, at least, moral properties as response-dependent theorists conceive of them, i.e. as nonnatural properties - seem to be causally inert. And the fact that response-dependent theories lack the resources to show that moral properties figure in causal explanations, the argument would go, has an impact not only on the issue of whether the Euthyphronist view about the direction of dependence between property and response can be rejected (and hence, on the issue of whether it can plausibly be held that there is real indeterminacy in the world), but also on the issue of whether uniformity of response can be regarded as the result of a reliable epistemic mechanism. (Recall that in order to respond to the argument from moral disagreement, one needs to show both that some moral issues are indeterminate and that convergence in moral belief can be regarded as the result of a reliable epistemic mechanism). The argument developed in the next chapter touches on the aforementioned issues. I will argue that moral properties can play causal explanatory roles, and that, despite common assumptions to the contrary, this claim can consistently be endorsed by both ethical naturalists and ethical nonnaturalists.
# Chapter 3 The argument from explanatory impotence

# 3.1 Harman's argument

I turn now to the argument from explanatory impotence. The debate over whether reference to moral facts is needed to explain our moral beliefs was initiated by Gilbert Harman (1977: ch.1).<sup>1</sup> Harman asks us to consider the case of someone who sees a group of children setting a cat on fire and immediately thinks "That's wrong". The question is whether we are required to refer to the actual wrongness of the children's action in order to explain why the person involved forms the belief that what the children are doing is wrong. According to Harman, such an explanation can dispense entirely with any reference to the actual wrongness of the children's action. In order to explain the belief in question, the argument goes, we only need to assume certain *non-moral* facts about the situation being observed (such as the fact that the children are really pouring gasoline on the cat, thus causing the cat extreme and unnecessary suffering), and certain non-moral facts about our upbringing and psychology (such as the fact that we are schooled to find repugnant the infliction of unnecessary suffering). Consider, by contrast, the case of a physicist who sees a vapor trail in a cloud chamber and thinks "There goes a proton". Harman grants that in this case, as in the previous one, in order to explain the belief about the presence of a proton we need to make certain assumptions about the observer's psychology (e.g. about the scientific theory he accepts and about his beliefs regarding the employed experimental device). However, the crucial difference between the two cases is that we are required to mention the fact that there really is a proton going through the cloud chamber as part of a reasonable explanation of why the physicist forms the belief about the presence of a proton. More specifically, it is reasonable to assume that there is a proton in the cloud chamber because the passage of the proton would explain the vapor trail, which would explain the physicist's visual condition, which, in turn, together with the physicist's scientific training, would explain his belief about there being a proton.<sup>2</sup>

<sup>&</sup>lt;sup>1</sup> For more recent versions of the argument from explanatory impotence, see Williams (1985*a*: ch. 8), Gibbard (1990: ch. 6), Wright (1992: ch. 5), and Leiter (2001).

<sup>&</sup>lt;sup>2</sup> One line of attack against Harman's argument is to question the claim that, whereas the belief that the children's action is wrong can be explained solely by reference to one's background moral beliefs, the belief that there is a proton in the cloud chamber cannot similarly be explained by reference to one's background scientific beliefs. For this criticism of Harman's argument, see Wright (1992: 190–1) and Shafer-Landau (2003: 101–2). As Wright rightly points out, the physicist would have formed the belief about the presence of a proton in the cloud chamber even if the vapor trail had been caused by the presence of a different particle, so far unknown to us. If so, then it is not at all clear that the best explanation of why the physicist forms the belief in question should advert to protons. Wright generally casts doubt on the idea that the best explanation of why we hold any of our (defeasible) scientific theories should make reference to the states of affairs that purportedly confer truth upon them: "ought not the best explanation of why [a physical theory] is held to be true to be *the same* whether it is actually subsequently defeated or survives indefinitely? After all, the causal antecedents of our accepting it are *already in place*, as it were, whatever the fate of the theory subsequently proves to be." (1992: 190, emphases in original) However, in what follows I will sidestep this complication. For even if one's background scientific beliefs together with the observation of the vapor trail seem sufficient to explain the belief about the presence of a proton (just as

Given this sharp contrast between explanation in ethics and in science, Harman concludes that moral beliefs do not possess the epistemic credentials possessed by scientific beliefs.

Harman's argument is underwritten by what we may call the *Explanatory Criterion*. According to this criterion,

The only facts and properties we are justified in believing in are those that are required to explain (i.e. that figure in the *best explanation* of) why we form the beliefs that we do.<sup>3</sup>

In her discussion of Harman's skeptical challenge, Judith Jarvis Thomson (1996: 69–76) contributes a useful clarification of the basic idea behind the *Explanatory Criterion*. Thomson points out that the epistemological arguments for moral skepticism have traditionally relied on the claim that moral sentences are not entailed by factual sentences. By conjoining this claim with the claim that moral sentences are epistemologically parasitic upon factual sentences (i.e. our only way of coming to know that a moral sentence is true is by coming to know that certain factual sentences are true), it follows

<sup>3</sup> Harman's argument is an epistemological argument, i.e. it is not an argument to the conclusion that there are no moral facts, but an argument to the effect that we lack evidence that moral facts exist. Yet, the argument from explanatory impotence is often formulated as an argument for the ontological claim that there are no moral facts. This argument relies on a modified version of the Explanatory Criterion, which holds that the only facts and properties that exist are those that are required to explain (i.e. that figure in the best explanation of) why we form the beliefs that we do. Those who defend the view that at the core of the idea that morality is objective lies the epistemological thesis that we can find out about some moral sentences that they are true (i.e. we can have adequate evidence that they are true), seem to be right about the primacy of epistemological issues at least in the case of the argument from explanatory impotence. The reason why we may want to eliminate certain facts or properties from our ontology is precisely that we do not have adequate evidence that they exist. For this view, see Thomson (1996: 68, 78). (One can point out that the principle of parsimony is already relied upon at an earlier stage of the argument, as one of the very criteria used in determining what counts as the 'best explanation' of a given phenomenon. Thus, the pressure to simplify our ontology is already felt when deciding on the best explanation of our moral beliefs. However, it is not entirely clear that this fact tells against the primacy of the epistemological. Whether or not this is the case ultimately depends on whether going from the thought that certain facts figure in the best theoretical explanation of a given phenomenon to the thought that these facts exist needs to be mediated by the thought that the facts in question exist only insofar as there is evidence that they exist and their featuring in the best explanation of a given phenomenon constitutes such evidence.) Now, although the links between the epistemological and the ontological issues at hand are close enough for the Explanatory Criterion to be easily (re)formulated one way or the other, one should not lose sight of the fact that, unless it is coupled with further assumptions, the epistemological version of the Explanatory Criterion cannot be used to undercut the claim that there are moral facts.

The reason why I am mainly interested here in the epistemological version of the argument has to do with the fact that, in the previous section, I have argued that whether the cognitivist can successfully reply to the argument from disagreement depends in part on whether he can successfully reply to the argument from explanatory impotence. Hence, my interest is in whether (convergence of) moral belief can suitably be regarded as the result of a reliable epistemic mechanism. What matters for my present purposes is not so much whether there are moral facts, but rather whether moral beliefs are ever justified.

one's background moral beliefs together with the observation of the children's action are sufficient to explain the belief that what they are doing is wrong), there still seems to be a relevant difference between the moral and the scientific case. In the scientific case, a physical entity *causes* an observable phenomenon (the vapor trail), and it is on the basis of this observable phenomenon that we can infer the existence of the cause (whether it be a proton or a different particle). However, no similar causal relationship exists between the wrongness of the children's action and an observable phenomenon. One of the questions that I address in this subchapter is whether moral facts can figure in the explanation of any observable phenomena.

that there is no way of finding out that a moral sentence is true. However, one can reject this traditional version of the argument for moral skepticism by pointing out that it mistakenly assumes that we cannot come to know that a sentence is true unless the premises from which we drew it as a conclusion actually entail the sentence in question. Undoubtedly, this view about the conditions under which we can come to know that a sentence is true is too stringent. As Thomson rightly emphasizes, in some cases our coming to believe that a sentence X is true conforms to the following pattern: we find out that certain sentences XX are true, and then we draw the conclusion X from them, where XX do not entail X, but the truth of XX nonetheless provides evidence for X. Thomson illustrates this point by means of an example. Suppose one finds that Smith has been shot, and John is standing next to Smith's body with a smoking gun in his hand. Why does one conclude that it was John who shot Smith? Although the sentence "Smith has been shot, and John is standing next to Smith's body with a smoking gun in his hand" does not entail that John shot Smith, its truth nevertheless provides evidence that he did. Now, surely, the truth of a sentence X' can count as evidence for the truth of a sentence X only insofar as there is some close connection between the truth of X' and the truth of X. What Harman (1965) proposes is in fact the following account of the above-mentioned connection: the truth of a sentence X' counts as evidence for the truth of a sentence Xonly if the truth of X would *explain* the truth of X'.<sup>4</sup>

Harman (1977) takes the matter a stage further and claims that ethical belief compares unfavorably with scientific belief: sentences reporting ethical belief cannot be explained by ethical truth in the way that sentences reporting scientific belief can be explained by scientific truth. Thus, in Harman's example, the physicist's belief that there is a proton passing through the cloud chamber counts as evidence for the existence of the proton because the physicist's belief is explained by the fact that there actually is a proton passing through the cloud chamber. By contrast, the belief that the children are wrong to set the cat on fire cannot count as evidence for the fact that what they are doing is really wrong because the actual wrongness of the children's action is completely irrelevant to the explanation of why the person involved forms the belief in question. If we accept the *Explanatory Criterion* of justified belief, and furthermore, if Harman is right that no factual sentence can be evidence for a moral sentence, since there is no factual sentence whose truth would be explained by the truth of a moral sentence, then moral knowledge is threatened by skepticism.

Before going on to discuss the merits of Harman's argument, a further qualification is needed. Consider the following passage from Harman:

<sup>&</sup>lt;sup>4</sup> According to Harman (1965), all nondeductive (probable) inferences can be treated as inferences to the best explanation. Harman contests the Peircean distinction between *induction* and *abduction* (or *hypothesis*). He argues that, whereas there are inferences which cannot be described as instances of enumerative induction, all (apparent) cases of enumerative induction can be described as instances of the *inference to the best explanation* (Harman's name for *abduction*). He stresses that describing enumerative inductions as inferences to the best explanation helps to make clear the conditions under which inferences are warranted (i.e. the inference from "All the observed A's are B's" to "All A's are B's" is warranted just in case the hypothesis that all A's are B's is, in light of all the evidence, better than any hypothesis to the contrary). It is worth noting, however, that Harman's view that *all* nondemonstrative inferences to the best explanations is not entertained by other advocates of inference to the best explanation. See, for instance, Lipton (2004).

Facts about protons can affect what you observe, since a proton passing through the cloud chamber can cause a vapor 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. (Harman, 1977: 7–8)

Harman's stress on the idea that moral facts, as opposed to physical ones, exert no causal influence on our perceptual apparatus suggests that he takes the explanatory role that any facts must play if our belief in them is to be vindicated to be *causal* in nature. According to this interpretation, the criterion of justified belief on which Harman's skeptical challenge relies is different from the one formulated above. The *Explanatory Criterion* needs to be strengthened by requiring that real facts and properties figure as causes in the explanation of our forming the beliefs that we do. This stronger criterion, which I will henceforth refer to as the *Causal Criterion*, states that:

The only facts and properties we are justified in believing in are those that exert a *causal* influence on our perceptual apparatus.<sup>5</sup>

In other passages, however, Harman suggests that the *Causal Criterion* of justified belief is too strong.<sup>6</sup> For example, Harman (1977: 9–10) raises the question whether ethics should be compared with mathematics, instead of physics. A common criticism of the causal theory of knowledge is that it renders mathematical knowledge impossible. Insofar as some sort of causal chain relating the knower and the object known is a necessary condition of knowledge, it seems that we cannot have mathematical knowledge, since we cannot possibly interact with mathematical objects. This unwelcome consequence has led many to reject this theory, and the corresponding *Causal Criterion* of justified belief. Harman seems to share the view according to which the *Causal Criterion* should be relaxed in order to make room for mathematical knowledge. Yet, he contends that endorsing a weaker criterion is hardly helpful to the moral realist since ethics compares unfavorably not only with physics, but also with mathematics. Harman writes that:

In explaining the observations that support a physical theory, scientists typically appeal to mathematical principles. [...] 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. (Harman, 1977: 10)

In other words, moral truths, as opposed to mathematical ones, are not needed to explain our holding *any* of the beliefs that we do. So, according to Harman, moral facts would fail to meet not only the *Causal Criterion*, but also the *Explanatory Criterion*.

<sup>&</sup>lt;sup>5</sup> Here I draw on Sayre-McCord's reconstruction of Harman's argument. See Sayre-McCord (1988b).

<sup>&</sup>lt;sup>6</sup> For suggesting that the *Causal Criterion* of justified belief is too strong and that we should opt instead for the *Explanatory Criterion*, see also Harman (1973: 126–32).

## 3.2 Causal and explanatory reasoning

Obviously, the realist's reply to the argument from explanatory impotence will vary according to the specific criterion of justified belief that is thought to be more appropriate. However, a brief survey of the literature will reveal that the question concerning the correct criterion of justified belief is still under dispute. Those who argue that the *Causal Criterion* is too strong and that we should opt instead for the *Explanatory* Criterion are motivated by the worry that some of our main epistemological achievements (such as mathematical claims or the laws of physics) would fail to meet the former criterion. Yet, the worry that motivates the opposite train of thought, according to which the Explanatory Criterion is too weak, is not unfounded. To put it bluntly, the idea is that whether something is causally potent is an objective matter.<sup>7</sup> So, if certain facts are causally potent, then we are justified in believing that they exist. By contrast, being merely explanatorily useful should not be allowed to count as evidence for the existence of the explanans.<sup>8</sup> As Shafer-Landau puts it, "what counts as a good explanation will, in many contexts, depend crucially on our interests and epistemic limitations" (2003: 99). This is precisely the reason why some philosophers think it is best to raise the standards for what counts as evidence for our beliefs to include not only explanatory potency, but also causal potency. Let me detail.

It is worthwhile noticing that there are two distinct ideas in the passage quoted above. On the one hand, the fact that our *interests* play a crucial role in determining what counts as an appropriate explanation has a bearing on whether being explanatorily useful should be allowed to count as evidence for more than just the belief in the existence of the explanans, non-realistically construed. On the other hand, the fact that what qualifies as a good explanation depends on our *epistemic limitations* has nothing to do with whether the facts that figure in such explanations should be construed non-realistically rather than realistically.<sup>9</sup> The idea is rather that passing an explanatory test cannot count as evidence for believing in the existence of any of the facts that we commonly suppose are best construed realistically. Perhaps the best way to illustrate this difference is to point out that, in some cases, we may be confident that there is nothing wrong with an explanation, and still be unwilling to allow that being explanatorily helpful is more than just an evidence for the existence of the explanans, non-realistically construed. (For instance, one may have no complaint against the claim that the reason why we think Charlie Chaplin is funny is that he really is funny, as long as 'being funny' is construed non-realistically.) In other cases, however, supposing there was a way to ensure that no cognitive shortcomings intervene in deciding what is required to explain a given

<sup>&</sup>lt;sup>7</sup> For this way of framing the issue, see Shafer-Landau (2003: 99).

<sup>&</sup>lt;sup>8</sup> This claim needs to be further qualified to allow that being explanatorily useful can, in any event, count as evidence for the existence of the explanans, so long as it is construed non-realistically. However, one should not lose sight of the fact that it is generally agreed by both proponents and opponents of the argument from explanatory impotence that the crux of the matter is whether we are justified in believing in moral facts, realistically construed. Harman himself does not claim that there are no moral facts. He only claims that, on a nonrelativist construal, there are no moral facts. See Harman (1996).

<sup>&</sup>lt;sup>9</sup> In fact, it can be argued that the closer one moves towards the idea that certain facts are best construed non-realistically, the less sense it makes to bring into discussion our epistemic limitations. To be sure, there are standards of warranted assertibility that govern areas of discourse which are best construed non-realistically. For this line of argument, see Wright's (1992).

phenomenon, there would be no room left for doubting the existence of the explanans, *realistically* construed. (Think of the hypothesis that continents were once joined together into a single supercontinent and drifted apart to their current location. The existence of a supercontinent would explain several facts, such as the striking fact that the edges of continents seem to fit together like a giant jigsaw puzzle, or the fact that identical fossil plants and animals were found on separate continents. If there was a way to ensure that the existence of a single original landmass is the best explanation of the aforementioned facts, there would be no room left for doubting its existence. In fact, our seeking for an explanation that is couched in causal terms is intended to guarantee that we come as close as possible to the best explanation. In the absence of a detailed causal story about how the original supercontinent broke apart into continents that drifted to their present location, we can have no guarantee that the proposed hypothesis is correct). Thus, a causal test is supposed to help in telling apart real facts from facts whose existence is best construed non-realistically, as well as from facts that do not exist.

There are two points I want to make in relation to the idea that a causal criterion of justified belief is more appropriate than an explanatory criterion. First, placing too much emphasis on the claim that whether something is causally efficacious is an objective matter may mistakenly suggest that causal explanations are not vulnerable to the criticisms leveled at explanations that are not couched in causal terms. Let me elaborate.

Consider a common criticism of the so-called inference to the best explanation. Advocates of inference to the best explanation typically hold that explanatory considerations are a guide to truth. On this view, the fact that a hypothesis exhibits the relevant explanatory virtues (i.e. the fact that it fares better than any other hypothesis with respect to simplicity, consilience, lack of *ad hoc* features, and so forth) is a reliable indicator of its truth. Yet, opponents of inference to the best explanation doubt that there is any connection between explanatory virtues and truth.<sup>10</sup> Why should we think, for instance, that a hypothesis's being simpler than another indicates that the first hypothesis is more likely to be true? This worry is sometimes voiced by asking whether there is anything that entitles us to assume that nature is simple rather than complex and hence that simpler explanations are to be preferred to more complex ones. Furthermore, why should we think that our pursuit of consilience is conducive to truth? Although offering a unifying account of previously disparate phenomena is definitely a virtue of an explanation, it is far from obvious that this is an epistemic rather than a pragmatic virtue. To paraphrase Nancy Cartwright (1983: 13), given that our knowledge of nature is highly compartmentalized, what reason is there to think that nature itself is unified?<sup>11</sup> Unless such questions are suitably answered, it is not at all clear that explanatory virtues ought to be treated as epistemic virtues. Opponents of inference to the best explanation also take issue with the uniqueness claim implied by the comparative term *best*. The idea is that, given our pursuit of a number of distinct explanatory virtues, sometimes there seems to

<sup>&</sup>lt;sup>10</sup> For a survey of the attractions and liabilities of inference to the best explanation, see Lipton (2004).

<sup>&</sup>lt;sup>11</sup> Cartwright's view on unifying laws is best summarized by the title of her book *How the Laws of Physics Lie* (1983). Cartwright argues that the laws of physics fail to describe nature as a whole. According to her, nature is "best described by a vast array of phenomenological laws tailored to specific situations" (1983: 66). Thus, read literally as descriptions of nature, the fundamental laws of physics are false; read as fictionalized descriptions, they do, nevertheless, play an instrumental role.

be no way of identifying the *best explanation*. Although in some cases explanatory virtues reinforce each other, in other cases they are in tension. When different explanatory virtues pick out different explanations, which one is to be assigned more weight in determining the best explanation?

Here is where causal explanation comes into the picture. Restricting explanation to causal explanation is supposed to insure both that we do not accept a hypothesis merely on the basis of its *pragmatic* virtues (i.e. on the basis of reasons that are independent of questions of truth), and that a *unique* hypothesis qualifies as the best explanation of a given phenomenon. For instance, Cartwright (1983: 4-5, 9, 90-1) argues that, whereas in general explanatory success is no guarantee of truth, the required link between explanatory success and truth obtains in the special case of causal explanation. Her argument goes as follows. A causal explanation is successful only if the causal process described by it actually takes place. Yet, for an entity to cause something, it must actually exist. Therefore, accepting that a causal explanation is successful is tantamount to accepting that the described cause actually exists. It is precisely because causal explanation has an existential component to it which theoretical explanation lacks that truth is built into the notion of a successful causal explanation. The upshot of this argument is that, whereas inference to the best explanation is unwarranted, inference to the likeliest cause is a legitimate form of inference. Another way of arguing for the desirability of restricting explanation to causal explanation is to point out that for any given phenomenon there is a uniquely correct causal story to be told.<sup>12</sup> Focusing on causal explanation would have the advantage of avoiding the difficulties inherent in the attempt to pursue several, potentially conflicting, explanatory virtues and of ensuring that a single hypothesis qualifies as the best explanation of a given phenomenon.<sup>13</sup>

<sup>&</sup>lt;sup>12</sup> This formulation is meant to allow that there are numerous causes of a phenomenon which are all part of a single causal history. The idea is that for any causal sequence that is part of the causal history of a certain phenomenon, there is a uniquely correct causal story to be told. A natural objection to the causal model of explanation is that it is too weak. As Lipton (2004: 32) puts it vividly, the big bang is part of the causal history of an event is explanatory of only a few of them. In other words, most information about the causal history of an event is explanatorily irrelevant. Lipton's notion of *contrastive* explanatory in a particular context, whereas another piece of information is not. Lipton points out that our requests for explanatory information often take a contrastive form, i.e. we do not simply ask "Why P?", but "Why P rather than Q?" Once we formulate the request for explaning a certain phenomenon in contrastive terms, the relevant part of its causal history is dramatically narrowed down.

<sup>&</sup>lt;sup>13</sup> Cartwright (1983: 74–5) stresses that the requirement that there be a single best explanation is a requirement that ought to be self-assumed by advocates of inference to the best explanation. This form of inference derives part of its intuitive appeal from the idea that it would be an absurd coincidence for a variety of phenomena to be all explained by a certain law and not be, in fact, consequences of that law. Cartwright acknowledges that the above-mentioned idea has considerable appeal. Yet, she points out that it speaks in favor of our inferences to the best explanation only insofar as they satisfy a requirement of non-redundancy. The claim that it would be an absurd coincidence for different phenomena to be explained by a particular law and not be, in fact, consequences of that law, makes sense only if there are no alternative explanations which account for the phenomena in question in an equally satisfactory manner. Cartwright goes on to argue that, whereas theoretical explanations do not satisfy the requirement of non-redundancy, the opposite is the case with causal explanations. She emphasizes that modern physics allows for a single correct causal story, while encouraging alternative theoretical treatments of the same phenomena. Cartwright points out that different purposes are better served by different theoretical models, and claims that questions concerning the truth of these theoretical models are simply misguided. See Cartwright (1983: 11-2, 85-6).

Now, let us briefly address the question of how causal reasoning differs in practice from reasoning that proceeds in explanatory terms. Peter Lipton's account of contrastive causal explanation may be especially relevant here. According to him, causal and explanatory reasoning are closely intertwined. Lipton (2004: 135-41) draws attention to three general aspects of our inferential practice. First, he stresses that we have the tendency to frame our requests for causes as requests for explanations. Although it might be thought that there is no relevant difference between asking a causal question and asking a why-question, the difference between these two types of question becomes significant in the specific context of *contrastive* questions. Lipton has repeatedly emphasized that most of our why-questions have an implicit contrastive structure. When we ask, for instance, why the leaves turn yellow in autumn, we may actually want to know why they turn yellow in autumn rather than in spring, or why they turn yellow rather than blue. So, our explanation requests usually take the form "Why P rather than Q?" instead of simply "Why P?" Accordingly, a good explanation must account for difference between the fact P and the foil Q. Now, Lipton points out that asking for an explanation of the difference between a fact and a foil is not the same as asking what caused the difference between them. To illustrate this point, he uses the well-known example of the nineteenth-century Hungarian physician who discovered the cause of childbed fever by investigating the question of why the maternal mortality rate in one of the clinics of the Viennese maternity hospital was considerably higher than the mortality rate in the other one of its clinics. Lipton argues that it was entirely natural for Semmelweis to ask for an explanation of the difference between mortality rates. However, to say that he was looking for a cause of the difference has a strange ring to it, since this would suggest some kind of interaction between the two cases. The upshot of all these considerations is that, when we are looking for a cause, we often find it more natural to ask a why-question rather than a causal question about our contrastive evidence, and hence, we conceive of potential causes as potential explanations.

Second, Lipton (2004: 136–7) argues that we often find ourselves reasoning in explanatory terms when we eliminate competing hypotheses. One way to decide between competing hypotheses is to look for additional evidence that would help us discriminate between them. What we do is to rule out causal candidates that fail to account for the additional data. Lipton claims that it is more natural to think of this failure to account for new evidential contrasts in explanatory rather than in causal terms. One of the reasons why the term 'account for' is better glossed as 'explain' rather than as 'cause' has to do with the fact that sometimes a factor shared by fact and foil can be a cause of the fact, but it will fail to explain the contrast between them. For instance, oxygen is a factor that is present both in the case of the match that lights and in the case of the match that fails to light because it is wet, and while oxygen is a cause of the dry match lighting, it fails to explain the difference between the two cases. Lipton emphasizes that we often rule out a putative cause because of its failure to explain a new contrast, and not because of its failure to cause.

Third, Lipton (2004: 137) points out that our inferential practice is helped by constructing various causal scenarios and by considering to what extent they would account for the available data. More specifically, what we do is to ask what causes would account for the effects we observe and what effects the putative causes would have. Again, Lipton argues that this type of subjunctive thinking is more natural if we think in

explanatory terms rather than directly about causes. He stresses that, when we think about potential causes and effects, "we [...] find it convenient to think in explanatory rather than in bluff causal terms, because it is often easier to say what a factor would explain than it is to say what it would cause" (ibid). Lipton further emphasizes that the question whether a putative cause would explain a contrast has a clear answer even in cases when we are uncertain about whether it would cause the effect. Going back to example of childbed fever, can one exclude the possibility that epidemic influences were part of the causal history of the deaths in the clinic which had a higher mortality rate? Maybe such influences constitute a necessary condition for any case of childbed fever. Yet, what is clear is that the epidemic hypothesis could not explain the difference in mortality rates between the two clinics.

In short, Lipton's view is that explanatory considerations serve as a guide to causal inference. As he summarizes his own view on contrastive inference,

[A] criterion we use to evaluate the quality of potential explanations is the same as one we use to infer causes. By inferring something that would provide a good explanation of the contrast if it were a cause, we are led to infer something that is likely to be a cause. (Lipton, 2004: 73)

Yet, once we allow that inference from effect to cause requires an appeal to explanatory considerations, we may have to conclude that the criticisms leveled at inference to the best explanation apply equally to inference to the likeliest cause. Insofar as causal inference is run on explanatory grounds, there seems to be no guarantee that the inferential procedure will take us towards the discovery of truth.

Presumably, this claim would be endorsed by those who reject inference to the best explanation and argue that only inference to the most likely cause is legitimate. For example, Cartwright (1983: 6–7, 98) contends that causal inferences are warranted only under special circumstances. In her view, it is only seldom that we are in a position to legitimately infer a cause outside the conditions of a controlled experiment. Cartwright insists that, if inference to the likeliest cause is to be the cornerstone of our epistemology, special attention should be given to what makes a cause "likely". According to her, the fact that a causal hypothesis exhibits certain explanatory virtues does not in any way count towards establishing the truth of the hypothesis in question.<sup>14</sup> In order to accept a causal hypothesis, we must have good reasons to believe that other causal candidates can be ruled out and that the cited cause constitutes the only practical possibility. Direct experimental testing, where we manipulate what we believe is the cause of a phenomenon and check whether the effects change in an appropriate manner, is the surest, if not the only way to isolate the actual causes of a phenomenon.

Lipton (2004: 132–4) acknowledges that there is a gap between causal and explanatory reasoning. He stresses that, whereas thinking in explanatory terms is a self-conscious thought process, there are cases when we make a causal inference without giving it too much thought. When I flip the switch and the light goes on, I have no doubt that it is my flipping the switch that causes the light to go on. My confidence stems from

<sup>&</sup>lt;sup>14</sup> In fact, Cartwright makes an even stronger claim. She maintains that the pursuit of explanatory virtues takes us away from the truth. She claims that "there is no reason to think that the principles that best organize much will be true, nor that the principles that are true will organize much" (Cartwright, 1983: 53).

the fact that I can control the light with the flipping of the switch, so there is no need for me to consider explanatory issues. Lipton describes such cases as near-automatic applications of Mill's Method of Difference and points out that they give us reason to conclude that it is not the case that we always rely upon explanatory considerations in drawing causal inferences. Yet, his sophisticated analysis of our inferential practice shows that explanatory considerations are usually at work in the inferential play. To stick to Mill's Method of Difference, Lipton (2004: 19, 126-8) highlights that this method sanctions an inference to the causal status of a difference between the antecedents of a case where the effect occurs and a case where it does not, but it says nothing about cases where we infer the *existence* of a difference and not just its causal role. He argues that inference to the best explanation provides a natural description of this sort of cases where we infer the existence of a difference because it would explain a contrast. Lipton points out that the Method of Difference not only fails to apply to inferences of unobservable causes, but also tells us nothing about how we are led to consider observable but as yet unobserved differences. So applying the Method of Difference requires a broader inductive framework than the one supplied by Mill's very method, and inference to the best explanation is well suited to describe this broader framework.

Lipton rejects the notion of inference to the best explanation which is associated with an extremely passive picture of scientific inquiry, "where we take whatever data happen to be at hand and infer an explanation, and where the central judgment we must make in this process is which of a battery of explanations of the same data would, if true, provide the loveliest explanation" (2004: 90). He points out that a realistic picture of scientific inquiry must allow for a feedback between the process of hypothesis generation and the process of data gathering (2004: 83). On this view, explanatory reasoning plays a complex role both in the process of hypotheses. Let me detail.

According to Lipton (2004: 65, 82), one of the virtues of the deductive-nomological model of explanation is that it makes clear that it is a mistake to suppose that our inferential practice conforms to a pattern according to which we first gather all the relevant data and only then we go on to consider the hypotheses which apply to them. This is because, in order to determine the evidence that is relevant to a certain hypothesis, we need to ask first what would have to be the case if the hypothesis in question were true. Now, recall that Lipton holds that explanatory reasoning suits best this sort of subjunctive reasoning. So, it is by considering what a given hypothesis would explain, if it were true, that we determine what further evidence is relevant to it. On this view, explanatory reasoning plays a central role in directing inquiry. Once we determine what as yet unobserved contrasts a given hypothesis would explain, we have a clearer idea about what manipulations or controlled experiments we need to perform in order to decide whether the hypothesis in question is correct.

We are now in a position to better appreciate the role of explanatory considerations in guiding causal inference. When we are interested in finding out what causes a certain phenomenon, asking a why-question about our contrastive data focuses inquiry. Most causal hypotheses are non-starters, and asking whether a hypothesis would explain our contrastive evidence helps narrowing down the range of candidate hypotheses that are worth considering. Furthermore, as we have already seen, reasoning about what would have to be the case if a certain hypothesis were true plays an essential role in directing observation and experiment. As Lipton describes the process, "contrastive data suggest explanatory hypotheses, and these hypotheses in turn suggest manipulations and controlled experiments that may reveal new contrasts that help to determine which of the candidates is the best explanation" (2004: 83).<sup>15</sup> Often it is on the basis of a process of manipulation or a controlled experiment that we are able to infer the truth of a certain hypothesis. However, this is not always possible, either because the sort of controlled experiment that would be needed to verify a certain hypothesis cannot possibly be set up, or because the processes of manipulation that we are able to perform leave more than one candidate hypotheses in the running. Lipton points out that when the experiments we perform do not succeed in eliminating all but one of the considered hypotheses, we resort to various explanatory considerations (e.g. considerations of scope, simplicity or consilience) in order to decide which hypothesis is correct. Yet, it is important to note that explanatory considerations come into play even when the processes of manipulation that we perform disconfirm all but one of the candidate hypotheses.

Consider again the example of Semmelweis's research into the causes of childbed fever. Semmelweis focused his inquiry by asking what would explain the difference in mortality rates between the two clinics of the Viennese maternity. He rejected several hypotheses on the grounds that they failed to explain this contrast (i.e. the hypotheses that childbed fever is caused by epidemic influences, overcrowding, diet or general care), and looked for additional evidence that would help discriminate between those hypotheses that did provide a potential explanation of the contrast (i.e. the hypotheses according to

<sup>&</sup>lt;sup>15</sup> One line of attack against Lipton's guiding claim is to argue that explanatory considerations play no role in the process of hypothesis generation. According to this line of argument, the fact that a causal hypothesis makes it onto the limited list of candidate hypotheses from which we subsequently select a single hypothesis has to do with its being reasonably plausible, where judgments of plausibility are independent of explanatory considerations. Yet, Lipton (2004: 149-51) points out that once we bring into discussion the inferential role of background beliefs, it becomes clear that the generation of a pool of plausible hypotheses rests on explanatory grounds. There are two points that are worth mentioning here. First, given that our background beliefs play a crucial role in restricting the range of plausible hypotheses, if background beliefs are based on explanatory considerations, then so are the judgments of plausibility that they generate. Second, the fact that the mechanism by which candidate hypotheses are generated favors the ones that cohere with our background beliefs testifies to the role played by the explanatory virtue of unification in the process of hypothesis formation. Our tendency to consider only hypotheses that can be seen as extensions of already accepted explanations helps produce a unified explanatory scheme. (According to Lipton (2004: 139), unification is a broad concept incorporating considerations of simplicity and consilience. The fact that we generally tend not to consider the hypotheses which conflict with many of our background beliefs promotes both simplicity and consilience.)

Clarifying exactly in what sense hypothesis generation relies on explanatory considerations goes some way towards meeting an objection that can be addressed to the argument presented here. One may concede that Lipton is right in claiming we often find it more natural to think in explanatory terms rather than directly about causes, while insisting that the sense of 'explanatory' that is relevant here is different from the one that Cartwright, for instance, finds objectionable. Cartwright has no intention of denying that reasoning about what explains what in a deductive or derivational sense plays an important role in scientific inquiry. (Clearly, there is no inconsistency between denying that all causal relations can be captured by derivational relations, a position that Cartwright would most certainly entertain, and admitting that deductive reasoning plays an essential role in many cases of causal scientific explanation). Nor would she deny that the process of conjecture is commonly conducted by asking which causal hypotheses would *explain* our contrastive data. Cartwright objects to the idea that the initial plausibility of a causal hypothesis, or the fact that such a hypothesis is ultimately deemed to be correct, has anything to do with whether the hypothesis in question exhibits certain explanatory virtues.

which the difference in mortality rates was due to a difference in birth positions, or to a difference in exposure to the priest, or to an infection by cadaveric matter). Semmelweis was led to the cadaveric hypothesis by his prior conjecture that the difference in mortality rates was somehow to be explained by the fact that women in the clinic with the higher mortality rate were attended by medical students, while women in the other clinic were attended my midwives. Given that medical students performed autopsies before performing examinations, the cadaveric hypothesis would explain why the mortality rate was lower both among women who delivered outside the hospital and among women who were attended by midwives. In order to test this hypothesis, Semmelweis insisted on students washing their hands with disinfectant after performing autopsies and found that the mortality rate in the clinic where students trained went down dramatically. At the same time. Semmelweis found that the mortality rates were not affected by a change in birth position or by the fact that the priest stopped visiting the hospital. So Semmelweis inferred that childbed fever was caused by an infection with cadaveric matter. Yet, Lipton points out that whereas the cadaveric hypothesis was successful in explaining the initial contrast in mortality rates between the two clinics, as well as the additional data revealed by the processes of manipulation performed by Semmelweis, it failed to explain other contrasts that were part of the evidence considered by Semmelweis. For instance, it failed to explain why some women who were attended by midwives contracted childbed fever while others did not, or why some women who delivered outside the hospital contracted the disease. Lipton (2004: 79) argues that, insofar as Semmelweis had good reason to believe that childbed fever was caused by an infection with cadaveric matter, this was only because he judged that the incompleteness of the cadaveric hypothesis constituted a better explanation of these explanatory failures than its falsity. So, explanatory considerations had a role in the inferential play even after Semmelweis performed the disinfection experiment. As Semmelweis's research shows, it is often difficult to make sense of an inferential decision without supposing that a choice is made between two overarching explanations, one of which states that a hypothesis's failure to explain a contrast is due to incompleteness, whereas the other states that explanatory failure is due to incorrectness.

The upshot of all these remarks is that we typically rely on explanatory considerations in making causal inferences. Yet, insofar as Lipton is right in claiming that inference to the best explanation should not be seen as in competition with causal inference, but rather as a way of realizing the latter, the criticisms that are commonly leveled at inference to the best explanation apply equally to inference to the likeliest cause. In other words, if the worry that the appeal to explanatory considerations makes the success of our inferences improbable is well founded, then we have reason to doubt the reliability of our causal inferences.

However, one can argue that none of the considerations adduced here proves that we are wrong in thinking that a causal criterion of justified belief is to be preferred to an explanatory criterion. One can stress that, while Lipton holds that explanatory considerations serve as an important guide to causal inference, he acknowledges that such considerations are not the sole guides to inference. As we have already seen, explanatory considerations play an important role in directing experiment, as well as in selecting from among causal hypotheses when the experiments we perform leave more than one hypothesis in the running. Still, the argument would go, we should not neglect the crucial role that experiments play in determining what inferences to draw. So even if explanatory considerations were not conducive to truth, this would not entail that we should abandon the causal criterion of justified belief, since explanatory considerations are not the only considerations that bear on causal inference. As Lipton (2004: 140) himself emphasizes, contrastive inference combines the advantages of both explanatory and causal reasoning, and this view is compatible with holding that a causal criterion of justified belief is more appropriate than an explanatory criterion.

Yet, I believe that Lipton's view that explanatory considerations are a guide to causal inference licenses a different line of argument. Insofar as we accept that Lipton's analysis of our inductive practices is descriptively adequate, and furthermore, insofar as we take these practices to be reliable, we will tend to discount as unfounded the worry that explanatory considerations are not truth-conducive. This is an argumentative strategy that Lipton himself pursues (2004: 144–8). He points out that inductive inferences, by their very nature, are not guaranteed to always yield the truth, and therefore, it is not an objection to explanatory reasoning that it does not make our inductive practices perfectly reliable. So there is no need to argue for a perfect overlap between explanatory and inferential virtues. Still, given that we generally take our inductive practices to be fairly reliable, and given that explanatory considerations are an important guide to inference, it would be quite surprising if such considerations were not truth-conducive. Hence, by showing that causal and explanatory reasoning are closely intertwined, Lipton helps to diffuse the objection that explanatory reasoning cannot reliably take us towards the truth.

This brings me to the second point I want to make in relation to the idea that the causal criterion of justified belief is more appropriate than the explanatory criterion. Our preferring a causal criterion to an explanatory one may have to do with a further observation made by Lipton. Consider the following two excerpts from Lipton:

[W]e often find it easier to reason in physical than in logical terms, in terms of causes rather than in terms of logical relations. [...] And one reason we find physical thinking so congenial is that our inductive reasoning is often abetted by processes of simulation, where we run causal scenarios in order to determine what inferences to draw. (Lipton, 2004: 132)

[W]e often think about inference in physical rather than logical terms, by constructing causal models and simulations rather than by investigating logical relations between statements or propositions. I take it that this thinking in the material mode is an aid to effective inference, but limiting one's attention to causal relations is restrictive, as compared to what is in a sense the universality of logical relations. My speculation is that by moving from thinking just about causation to thinking about explanation, we in effect compensate for this limitation, by being able to incorporate into our inferential thinking the very diverse range of relevant considerations [...]. (Lipton, 2004: 140)

On this view, one of the reasons why we find the causal model of explanation so natural is that thinking in physical terms is more congenial to us than thinking in strictly logical terms. Yet, even if we tend to think about inferential relations in physical terms, this does not mean that all explanation is causal explanation. Whether we should accept that there are adequate non-causal explanations remains an open question, especially since, as we have already seen, the worry that explanatory reasoning cannot take us towards the truth is ill founded.

By way of conclusion, let me say that none of the considerations presented in this section was intended to establish conclusively that the causal criterion of justified belief should be abandoned in favor of a weaker, explanatory criterion. My aim was to show that, according to one of the most sophisticated analyses of our actual inductive practices, causal inference requires an appeal to explanatory considerations, and that this helps to allay the suspicions regarding explanatory reasoning that motivate the rejection of the explanatory criterion in the first place.

Before turning to consider the realist's reply to the argument from explanatory impotence, it is perhaps also important to emphasize that authors like Cartwright reject the fundamental laws of physics precisely because they do not countenance inference to the best explanation. Cartwright's view is symptomatic of the difficulties that anyone who rejects the explanatory criterion of justified belief must deal with. If Cartwright is right in thinking that we should hold tight to a causal criterion, then many of our ordinary or scientific beliefs would be discarded. On this view, our moral beliefs are on an equal epistemic footing with some of our most entrenched scientific beliefs, a result that proponents of the argument from explanatory impotence would definitely want to avoid.

#### 3.3 Ethical naturalism

Let me turn now to the argument from explanatory impotence. As I have pointed out in section 3.1, certain passages from Harman suggest that, in his view, the causal criterion of justified belief is too strong, and that the argument from explanatory impotence must turn on the weaker, and more plausible, explanatory criterion. However, most of the ensuing literature focuses on the question whether moral facts and properties are causally inert. There are two main lines of response to the argument from explanatory impotence. Some moral realists accept the causal criterion and argue that moral facts are causally efficacious. Others claim that we have reason to reject the causal criterion, and therefore, it is no objection to moral realism that moral facts do not figure in causal explanations.

The first line of reasoning mentioned above is generally adopted by those who defend naturalistic versions of moral realism.<sup>16</sup> According to ethical naturalism, moral properties are "nothing over and above" natural properties. Proponents of reductive naturalism hold that moral properties are identical to natural properties that can be picked out in nonmoral terms. Proponents of non-reductive naturalism hold instead that moral properties are irreducible natural properties in their own right, which we may be unable to pick out in nonmoral terms. Given that ethical naturalists claim that moral facts are part of the natural fabric of the word, answering the challenge posed by the argument from explanatory argument is especially important to them. This is because proving that moral facts play an explanatory causal role provides a secure way of establishing that

<sup>&</sup>lt;sup>16</sup> This line of reasoning is adopted both by reductive naturalists (see Railton (1986; 1998)) and non-reductive naturalists (see Sturgeon (1985; 1986*a*; 1986*b*; 2006), Boyd (1988), and Brink (1989: 182–97)).

moral facts are on a par with other facts that are commonly viewed as part of the natural fabric of the world (such as physical, chemical or biological facts).<sup>17</sup>

Ethical nonnaturalists have generally found more congenial the second line of response mentioned above. According to this line of thought, even though moral facts do not figure in causal explanations, this is no objection to moral realism. The causal criterion of justified belief, which underlies the argument from explanatory impotence against moral realism, is rejected on various grounds. Some realists look for "companions in guilt". For instance, Russ Shafer-Landau (2006) and Terence Cuneo (2007*a*: 103–7) argue that epistemic facts (i.e. facts that concern what one ought to believe) are similar to moral facts in that they fail to satisfy the constraints of the causal criterion. So, the causal criterion would also give us reason to be skeptical about the existence of epistemic facts.<sup>18</sup> Given that endorsing this criterion would lead to such theoretically problematic consequences, Shafer-Landau and Cuneo conclude that we would do better to reject it. In a similar vein, Thomas Nagel emphasizes that the causal criterion would eliminate all normative truths and that this is a highly implausible result. He points out that:

[I]t begs the question to assume that this sort of explanatory necessity is the test of reality for values. The claim that certain reasons exist is a normative claim, not a claim about the best causal explanation of anything. To assume that only what has to be included in the best causal theory of the world is real is to assume that there are no irreducibly normative truths. (Nagel, 1986: 144)

Other moral realists point out that the existence of some sort of causal connection between the knower and the object known is what distinguishes empirical from *a priori* forms of knowledge, and contend that endorsing the causal criterion begs the question against *a priori* knowledge. For instance, Nick Zangwill (2006) argues that we should be careful not to build into the terms that define the debate over moral realism a prejudice against *a priori* knowledge, especially since, as he attempts to prove, there may be other, independent, reasons for thinking that moral knowledge is *a priori*.

<sup>&</sup>lt;sup>17</sup> According to Sturgeon (2006: 256, n6), moral realists who deny that moral facts are causally efficacious adopt this stance precisely because they believe that the implied ethical naturalism is implausible. However, it is important to keep in mind that the connection between being an ethical naturalist and holding that moral facts figure in causal explanations, on the one hand, and being an ethical non-naturalist and holding that moral facts do not play an explanatory causal role, on the other hand, is not as straightforward as it may first seem. Some ethical naturalists argue that whether moral facts play an explanatory causal role is irrelevant, since the main role of moral facts is not to explain action, but to justify it (see Sayre-McCord (1988*b*)), whereas some proponents of the so-called "neo-Aristotelian ethical naturalism" argue that moral facts are causally inert (see Thomson (1996: 67–94)). On the other hand, there are ethical non-naturalists who hold that moral facts can figure in causal explanations (see Wiggins (1990–91; 1996)).

<sup>&</sup>lt;sup>18</sup> Another way to express this worry is to point out there is something self-defeating about endorsing the causal criterion. According to this criterion, the only facts we are justified in believing in are those that figure in the best explanations of our experiences. However, the fact that an explanation is better than another is itself an evaluative fact that does not figure in the best explanation of our experience. The causal criterion would thus recommend its own rejection. For this line of argument, see Sayre-McCord (1988*b*: 277–9) and Shafer-Landau (2003: 113–5; 2006). It should be noticed that Sayre-McCord's criticism envisages the explanatory criterion. In what follows, I will not dwell on the question whether some of the arguments advanced by nonnaturalists would also apply to the explanatory criterion, but focus instead on arguing that nonnaturalists can adopt a different strategy, i.e. claiming that certain moral facts can play a causal explanatory role.

In what follows, I will attempt to prove that moral realists can successfully respond to the argument from explanatory impotence. I will start by considering the view that moral facts figure ineliminably in causal explanations. My aim will be twofold: first, to show that ethical naturalists are right in claiming that moral facts can play genuine causal explanatory roles, and second, to show that, despite what is commonly assumed, this argumentative strategy is available not only to ethical naturalists, but also to ethical nonnaturalists. I will then go on to consider the nonnaturalist view according to which we have reason to believe that the causal criterion of justified belief is too strong. Recall from section 3.1 that, according to Harman, moral facts would fail to meet not only the causal criterion, but also the explanatory criterion. The last two sections of this chapter are devoted to a discussion of David Wiggins's account of how moral facts can explain (non-causally) our beliefs about them. The picture that will emerge at the end of this chapter is one according to which the most defensible response to the argument from explanatory impotence involves two claims: first, the claim that some moral facts are causally efficacious, and second, the claim that moral facts figure ineliminably in the best explanation of our beliefs about them.

Before going on to discuss the realist's reply to the argument from explanatory impotence, let me explain why I think it is important to address the question of whether different forms of realism (i.e. naturalism and nonnaturalism) have the resources to respond to the challenge posed by this argument. Recall that in the previous chapter I have argued that the success of the realist's reply to the argument from disagreement is conditional upon her success in replying to the argument from explanatory impotence. So, deciding which version of realism fares better with respect to the latter argument might univocally determine which version of realism is well-suited to answer to the former argument. If it turns out, for instance, that naturalism is the only form of realism that has the resources to respond to the argument from explanatory impotence, then it would seem that, when we turn to the argument from disagreement, all nonnaturalistic versions of moral realism are automatically excluded. However, I believe that both naturalism and nonnaturalism have the resources to respond to the argument from explanatory impotence. The remaining of this chapter is devoted to defending this claim. More concretely, my aim is to show that a satisfactory response to the argument from explanatory impotence consists of a number of claims that can be embraced by both naturalists and nonnaturalists. Obviously, it is not my intention to deny that an elaborate treatment of certain issues connected with the discussion to follow will favor one version of moral realism over another. I am concerned to prove only that, when discussing the argument from disagreement, there is no reason to exclude from the outset some versions of moral realism (as it will become clear later on, I am mainly concerned with establishing that nonnaturalistic versions of moral realism should not be ruled out). The question whether naturalism is, on balance, a more plausible metaethical position than nonnaturalism lies beyond the scope of the present discussion.

I turn now to the view that moral facts are causally efficacious. One of the most ardent defenders of this view is Nicholas Sturgeon, a proponent of the so-called "Cornell realism".<sup>19</sup> Sturgeon (1985) argues that Harman was wrong in thinking that reference to moral facts is not needed to explain our moral beliefs. He draws our attention to an entire

<sup>&</sup>lt;sup>19</sup> The term "Cornell realism" is used to refer to non-reductive naturalism, as developed in the writings of Brink (1984; 1989), Sturgeon (1985; 1986*a*; 1986*b*; 2006) and Boyd (1988).

range of cases which seem to have gone unnoticed by Harman and which suggest that moral facts figure in the best explanation of our beliefs about them. To stick to one of the examples discussed by Sturgeon, the fact that Hitler was morally depraved forms part of a reasonable explanation of his deeds. In turn, Hitler's deeds form part of a reasonable explanation of our forming the belief that he was morally depraved. In short, Sturgeon claims that facts about one's moral character can plausibly be cited as part of a causal explanation of one's behavior and that these facts are then available as a reasonable further explanation of our arriving at a correct assessment of that person's moral character.

In what follows, I will consider two objections to the view that moral facts can figure in causal explanations. According to the first objection, some moral explanations are simply vacuous. For instance, Brian Leiter writes that:

My own feeling is that if I were seeking an explanation for Hitler's conduct and was offered the explanation "He was morally depraved", I would take such an answer to be a bit of a joke: a repetition of the datum rather than an explanation. (2001: 94, n53).

According to the second objection, although moral facts may seem to play causal explanatory roles, it is the nonmoral facts on which moral facts supervene that are in fact doing all the causal explanatory work.

Let me start by addressing the first objection. Many of us would probably agree with Leiter in thinking that Sturgeon's explanation of Hitler's deeds is vacuous. But why is it that Hitler's being morally depraved cannot provide an adequate explanation of his conduct? I believe that the problem with this sort of explanations has less to do with the fact that they are couched in moral terms, than with the unspecificity of the employed moral terms.

Consider another example used by Sturgeon (1985: 64), which concerns an incident that took place in the nineteenth century, when a group of American settlers recently arrived in California failed to save another group of travelers that were trapped by snow in the Sierra Nevada mountains. The rescue expedition was organized under the direction of Selim Woodworth. According to the sources cited by Sturgeon, although Woodworth volunteered to take charge of the rescue operation, he made absolutely no effort to carry out any of the tasks that might have led to the rescue of those trapped in the mountains, even tasks that required no heroism on his part. Sturgeon argues that, insofar as the evidence presented by historians is correct, it seems reasonable to cite the fact that Woodworth was "just no damned good" as part of an explanation of his deeds, and that this moral fact is available as a reasonable further explanation of our forming the opinion that he was no damned good. Sturgeon further emphasizes that "it would be difficult to find a serious work of biography [...] in which actions are not explained by appeal to moral character: sometimes by appeal to specific virtues and vices, but often enough also by appeal to a more general assessment" (ibid.). It is worth noticing that earlier on the same page Sturgeon suggests that Woodworth's deeds can be explained by appeal to specific vices. He writes that Woodworth "spent time arranging comforts for himself in camp, preening himself on the importance of his position; and as a predictable result of his cowardice and his exercises in vainglory, many died who might have been saved" (ibid., emphases added). However, he then chooses to formulate his conclusion in the language of thin moral concepts, thus inviting the objection that all his examples are patently vacuous.

There are two considerations that might have misled Sturgeon into thinking that explanations that are couched in thin moral terms can do a better job of explaining behavior than explanations that appeal to specific virtues or vices. First, Sturgeon's argument can go through if and only if a person's moral character forms part of the explanation of our arriving at a *correct* assessment of that person's moral character. In other words, it is not enough to prove that certain traits of character *can* form part of an explanation of a person's deeds; what needs to be proved is that competing explanations can be eliminated, i.e. that possessing those traits of character is the best explanation of why the person in question acted in a certain way. This would ensure that one can arrive at a correct assessment of that person's moral character. In what Woodworth's behavior is concerned. Sturgeon argues that all the evidence points against other explanations that compete with the one he favors. For instance, it seems hardly plausible that Woodworth was a basically good person who proved too weak in a situation in which a heroic gesture was expected of him. Sturgeon has purposely chosen examples that concern actions about which we would readily agree that they are wrong (in fact, we do not even have to reflect about them in order to conclude that they are wrong). However, it is not at all clear that this consensus would extend beyond judging such courses of action to be wrong. In particular, asking how an agent has come to act in a certain way and whether some specific flaws of character have played a role in this process would give way to differences of opinion, and hence, would weaken Sturgeon's case in favor of the idea that facts about an agent's moral character are part of the best explanation of how we come to believe in them. By contrast, keeping the discussion at a high level of generality creates the illusion that our consensus extends far beyond the claim that certain courses of action are wrong. When we are asked to consider Woodworth's behavior, we seem to agree with Sturgeon that only a person who was no damned good could have done something that is so obviously morally wrong. Likewise, when asked to consider Hitler's deeds, we do think that only a morally depraved person could have done something morally so abominable. And we seem to agree about such general assessments, regardless of whether we agree about what specific traits of character might be part of an explanation of how an agent came to behave in a certain way. Be that as it may, the sort of moral facts invoked by Sturgeon cannot figure in a causal explanation of an agent's behavior. The fact of being morally depraved cannot possibly explain Hitler's deeds, nor can such a fact explain any other instance of human behavior. This is simply because terms like 'morally depraved' lack a determinate descriptive content. So, it is as if Sturgeon was so intent on showing that the moral explanation he proposes cannot plausibly be replaced by competing explanations, that he failed to see that explaining behavior in terms of thin ethical concepts creates more problems than it solves. However, if facts about moral character are to figure in causal explanations of behavior, these can only be facts of the sort conveyed by thick ethical concepts.

A second consideration that could have led Sturgeon to think that explanations that are couched in thin ethical concepts can do a better job of explaining behavior than explanations that are couched in thick ethical concepts may have to do precisely with the fact that thick concepts have a heavy descriptive content. Thus, explaining behavior in terms of thick concepts may easily invite the objection that it is the underlying nonmoral facts (i.e. psychological facts) that are actually doing all the causal explanatory work. I will deal with this objection latter on. For now, what is important to keep in mind is that the only strategy available to advocates of moral explanations of behavior is to appeal to thick ethical concepts.

One can argue that even by adopting the aforementioned strategy, Leiter's objection cannot be avoided. According to Leiter, in order to see that Sturgeon's explanation of Hitler's deeds is vacuous, all we need to do is to compare this explanation to a finely nuanced psychological explanation of the sort provided in Erik Erikson (1963). Yet, is there sufficient reason to conclude that such psychological or psychoanalytic explanations of behavior are to be preferred to moral explanations that are cashed out in terms of specific moral vices? This is too big a question to properly address within the confines of this dissertation, but I will offer some considerations which suggest that the answer to this question is negative.

Why is it that we tend to agree with Leiter in thinking that, if we were looking for an explanation of Hitler's behavior, we would take Sturgeon's answer to be a bit of a ioke? The hidden appeal of Leiter's argument may have to do with the fact that when we ask what explains Hitler's deeds we are in fact more interested in a different question namely, how is it possible that something morally as abominable as the Holocaust has happened? This would explain why we tend to dismiss Sturgeon's explanation as unserious. Yet, if Hitler's moral traits cannot possibly explain the Holocaust, nor can his psychological traits. Leiter asks us to compare Sturgeon's moral explanation to a psychological explanation of the sort offered by Erikson. However, it is not at all clear that this comparison works to the credit of Leiter's argument. It is worthwhile noticing that Erikson aims at providing not only a psychological portrait of Hitler, but also a richly nuanced account of the psychological characteristics of an entire generation that supported the national socialist movement. Insofar as Erikson's account tells us something about the conditions that made the occurrence of the Holocaust possible, it does so by pointing to more than just Hitler's psychological traits. Likewise, if we were offered a moral explanation of why the Holocaust occurred, an explanation that referred to both Hitler and those who supported him (or those who failed to oppose him), we would no longer think that such an explanation is inadequate (although it remains the case that, if such explanations were cashed out in thin moral terms, we would rightly feel that they cannot be taken seriously).

Still, one might claim that, if we leave aside questions concerning the conditions that made the occurrence of the Holocaust possible and focus only on how to explain Hitler's behavior, it becomes manifest that psychoanalytic explanations provide better explanations of behavior than moral explanations. On this line of thought, the fact that Hitler instigated such atrocities is to be explained by reference to a pathological condition, rather than by reference to ordinary vices. However, even if we accept that Hitler could not have done what he did unless he suffered from a psychological disorder, at the end of the day this case does more to convince us that psychoanalytic explanations of behavior cannot always be accepted. Once we admit that the Holocaust could not have taken place without the complicity (or the passivity) of so many people, most of us will resist explaining the Nazi atrocities in psychoanalytic terms. This tendency to reject psychoanalytic explanations has to do with the fact that we strongly believe that we cannot exonerate all those people who supported in one way or another or actively contributed to Nazi atrocities. Historians of the Holocaust point out that, even if it can be argued that Hitler suffered from a psychological disorder, we should not lose sight of the fact that most of the people who contributed to Nazi crimes were quite ordinary people. The point is that it would be inappropriate to explain their behavior by reference to psychopathological features. Yet, our feeling that psychoanalytic explanations shed no light on why such atrocities have occurred goes deeper than this. This is a context in which the feeling that appealing to reductionist explanations of the sort offered by psychoanalysis is inappropriate is particularly strong. To paraphrase Leiter (2001: 87), such naturalistic accounts explain moral behavior by appeal to the *deterministic forces* operative in psychology (for instance, by establishing certain links between adult behavior and childhood experiences). Moreover, those who defend reductionist explanations of moral behavior claim that such explanations leave nothing more to explain (i.e. nothing that would require an appeal to moral facts). And this is precisely what we find so hard to accept when we reflect on what explains the occurrence of the Holocaust. In this context, most of us cannot help the feeling that psychoanalytic explanations are superficial, and can at best cover only a small part of the ground. This concludes my discussion of Leiter's objection against moral explanations.

Now, let me consider the second objection to the view that moral facts are causally efficacious. According to this objection, although moral facts may seem to play causal explanatory roles, it is the underlying nonmoral facts that are actually doing all the causal explanatory work.<sup>20</sup>

It should be clear from the outset that there is one version of moral realism that is not vulnerable to this objection. According to reductive naturalism, moral properties are either identical to, or else reducible to natural properties. So, if it turned out that reductive naturalism is the most defensible metaethical position, there would be no ground for doubting that moral properties play causal explanatory roles. The question is whether the above-mentioned objection can be avoided by those versions of realism according to which moral properties are non-reductively supervenient on natural properties.

Before addressing this question, let me first introduce the notion of supervenience. "Supervenience" refers to a relation of necessary covariance among properties. To claim that moral properties supervene on natural properties is to claim that no two things, whether acts or persons, can differ with respect to their moral properties unless they also differ with respect to their natural properties (or, equivalently, one can say that no two things can share all their natural properties without also sharing their moral properties). There is however a different relation that also goes under the name of supervenience. It is widely assumed that moral properties necessarily depend on moral properties. In other words, it is assumed that if something possesses certain moral properties this can only be "because of", or "in virtue of" its possessing certain natural properties. Although the idea that the moral covaries with the natural and the idea that the moral depends on the natural will be dealt with separately in the next section, for the time being I will just assume that they are part of a single principle.

The realist who holds that moral properties are non-reductively supervenient on natural properties faces the following problem. According to the supervenience principle, moral properties hold only in virtue of natural properties. So, when we have a case in which a moral property seems to play a causal explanatory role, there must be some

<sup>&</sup>lt;sup>20</sup> For this line of argument, see Harman (1986), Thomson (1996), Audi (1997), and Leiter (2001).

(possibly complex) natural property the instantiation of which is sufficient for the instantiation of the moral property in question. But then it seems that all the causal explanatory work is in fact done by the natural properties on which moral properties supervene, whereas moral properties are causally and explanatorily irrelevant.

A common response to this objection is that, if the fact that moral properties supervene on natural properties gave us reason to deny that moral properties possess causal explanatory powers, then we would have to conclude that chemical, biological or psychological properties, which supervene on physical properties, also lack such explanatory powers.<sup>21</sup> In other words, this objection would prove too much. Another way to respond is to attempt to show that moral explanatory loss.<sup>22</sup> For instance, Peter Railton argues that an explanation that refers to honesty, a normatively shaped category, conveys distinctive explanatory information that may be lost if we were to replace this explanation with one that referred only to the psychological dynamics on which honesty supervenes. He writes that:

Honesty amounts to different things in different contexts, and the interests to which it answers, though familiar, are quite complex features of the social landscape. Consider what honesty requires of Bob if: he is describing an illness to a child, or to an adult of great emotional vulnerability; or he has promised confidentiality to one friend but not another; or he is playing a friendly game of poker; etc. Let us suppose there to be no compact, otherwise psychologically unified set of characteristics that make up honesty. [...] We nonetheless think we are giving a distinctive kind of explanation when we attribute Bob's having come forward with self-incriminating evidence to his honesty, rather than, say, to his naiveté, or to a cooly self-interested calculation. (Railton, 1998: 179–180)

Answering to the charge of explanatory impotence is especially important to Cornell realists. This is because they hold that moral properties are on a par with other properties that are commonly viewed as part of the natural fabric of the world, and proving that moral properties pull their weight in explanatory theories seems to be the best way of showing that this is so. Note, however, that none of the considerations I have presented so far suggests that ethical nonnaturalism is incompatible with holding that moral properties can play causal explanatory roles.<sup>23</sup> Now, let us consider Railton's suggestion to the effect that there is such an incompatibility. Railton contends that:

Indeed, even outright reducibility is no ground for doubting explanatoriness. To establish a relation of reduction between, for example, a chemical phenomenon such as valence and a physical model of the atom does nothing to suggest that there is no

<sup>&</sup>lt;sup>21</sup> See, for instance, Railton (1986: 183; 1998: 178– 9), Sturgeon (2006: 249–50), and Shafer-Landau (2006).

<sup>&</sup>lt;sup>22</sup> See, for instance, Sayre-McCord (1988b: 275-6).

 $<sup>^{23}</sup>$  The view that moral properties are causally inefficacious is common among nonnaturalists. As Sturgeon rightly points out, most nonnaturalists adopt this view because they believe that to hold otherwise would mean to accept that moral properties are natural properties. See Sturgeon (2006: 256, n6). In the next section, I will argue against the assumption that holding that moral properties are causally efficacious and holding that moral facts are natural properties are intimately linked. For an argument to the same effect, see Cuneo (2007*b*).

such thing as valence, or that generalizations involving valence cannot support explanations. There can be no issue here of ontological economy or eschewing unnecessary entities, as might be the case if valence were held to be something *sui generis*, over and above any constellation of physical properties. The facts described in principles of chemical valence are genuine, and permit a powerful and explanatory systematization of chemical combination; the existence of a successful reduction to atomic physics only bolsters these claims. (Railton, 1986: 183–4)

In short, Railton claims that whereas the objection that supervening properties lack genuine explanatory power poses no problem to non-reductive or reductive naturalists, nonnaturalists, who hold that moral properties are something "over and above any constellation of physical properties", have trouble answering this objection. In the following section, I will attempt to prove that ethical nonnaturalism has the resources to respond to the aforementioned charge.

## 3.4 Moral supervenience

In this section, I will argue that the realist can resist the conclusion that moral properties are causally and explanatorily irrelevant, and that the real causal explanatory work is done by the natural properties on which moral properties supervene, by expressing doubts about the truth of the supervenience principle.

Recall that, in the previous section, I have considered the suggestion that moral traits of character are the most plausible candidates among moral facts for being causally efficacious. I have argued that, if facts about moral character are to figure in causal explanations of behavior, these can only be facts of the sort conveyed by thick ethical concepts. However, there seems to be a problem with appealing both to the notion of supervenience and to that of thick ethical concepts.<sup>24</sup> To claim that the moral supervenes on the natural is to claim that moral properties necessarily depend on (and covary with) natural properties. In other words, if an object possesses certain moral properties, this is because it possess certain natural properties. This principle presupposes a sharp distinction between the evaluative and the natural. Yet, once we consider thick ethical concepts, the aforementioned distinction is thrown into doubt. Arguably, in the case of thick ethical concepts, the descriptive and evaluative elements are indissolubly linked.<sup>25</sup> Someone who failed to grasp the evaluative point of a thick concept would not be able to predict its extension solely on the basis of descriptive similarities.<sup>26</sup>

<sup>&</sup>lt;sup>24</sup> For suggesting that the interconnectedness of the evaluative and the descriptive gives us reason to doubt the truth of the supervenience principle, see Griffin (1998: 44–6), Dancy (1996: 278–9). For expressing doubts about the truth of the supervenience principle, see also Raz (1999) and Sturgeon (2006). See also Cuneo (2007*b*: 361–2), for arguing that the realist can resist the conclusion that moral properties are causally inert by appealing to considerations about thick ethical concepts.

<sup>&</sup>lt;sup>25</sup> See Williams (1985*a*). Williams draws on McDowell (1981).

<sup>&</sup>lt;sup>26</sup> One can raise the following objection against the argument of this section. My aim is to show that the realist can win the debate over whether moral facts can figure in causal explanations by raising doubts about the truth of the supervenience principle. Yet, one can point out that insofar as this argument relies on the claim that the descriptive and the evaluative elements are indissolubly linked in a thick concept, this argument begs the question against noncognitivists, who typically defend the distinction between the

A brief terminological point is in order. I have been using here the terms "evaluative" and "descriptive". Discussion about thick moral concepts is usually phrased in these terms. However, the supervenience principle is more often expressed by saying that the moral depends on, and covaries with, the natural. Without implying that there is a perfect overlap between the descriptive-evaluative distinction and the natural-moral distinction, we can take as the claim that the advocates of moral supervenience intend to make the claim that the evaluative supervenes on the nonevaluative. Let me offer a brief indication of why the latter formulation seems more appropriate. According to ethical naturalism, moral properties are natural properties. So, as Sturgeon (2006: 248) points out, to say that moral or evaluative properties supervene on natural properties is to make an assertion that is true, but trivial. Sturgeon emphasizes that the advocates of the supervenience principle want to claim that evaluative properties necessarily depend on other properties, and that we need to have reasons for the ascription of evaluative properties; yet, none of these conclusions is implied by the aforementioned assertion. Thus, expressing the supervenience principle as the supervenience of the evaluative on the nonevaluative seems better suited to capture what the advocates of this principle have in mind. Although in what follows I will sometimes refer to the base properties on which moral properties supervene as natural properties, the term "natural" should be understood as meaning simply "nonevaluative".

Drawing attention to the interconnectedness of the evaluative and the descriptive is not the only way of arguing against the supervenience principle. One can also argue that the supervenience principle is more doubtful than it is commonly supposed to be by pointing out that there is no reason to think that there are enough nonevaluative predicates to cover all the existing nonevaluative properties. For instance, Sturgeon suggests that there are nonevaluative properties for which we do not possess nonevaluative terminology.<sup>27</sup> He emphasizes that physics supports the idea that there are more physical properties than there are predicates to represent them, since it appears that "there are continuum many physical states of the world, but there are at most countably many predicates in any language" (Sturgeon, 1985: 240). So, insofar as the supervenience principle is taken to mean that evaluative properties, i.e. properties represented by evaluative terms, supervene on properties that can be represented by nonevaluative terms, we have reason to doubt its truth.<sup>28</sup>

descriptive and the evaluative. While it is not possible to follow here all the intricacies of this debate, some of the considerations offered below are meant to show that the descriptive element of a thick concept is not independently comprehensible. Note also that whereas early noncognitivists held that ethical concepts that combine emotive or prescriptive meaning with descriptive meaning could be factored into two separate elements, other noncognitivists reject this view. See Williams (1985*a*). Moreover, modern expressivists attempt to accommodate the idea that evaluation partly determines description. See Blackburn (1992). For useful discussion, see Dancy (1996).

<sup>&</sup>lt;sup>27</sup> Sturgeon (1985) uses this argument to rebut the view that ethical naturalism must provide reductive naturalistic definitions for moral terms. Compare Sturgeon (2006: 248–9).

 $<sup>^{28}</sup>$  A similar idea is expressed by Raz. He writes that: "there is no reason to think that the evaluative predicates of English (or of any other natural language) supervene on its non-evaluative predicates. [...] In fact my doubt is more general. It concerns the truth of global evaluative supervenience theses in general. [...] Which predicates belong to a language at any given time is a contingent matter, a product of its historical development to that time, and so is the number and identity of the concepts available to any person, or any cultural group." (Raz, 1999: 220–1).

Now, let me consider in more detail how rejecting the supervenience principle would help the realist win the debate about whether moral properties are causally efficacious. As we have seen in the previous section, whenever we have a case in which a moral property seems to play a causal explanatory role, one can argue that all the causal explanatory work is in fact done by the (complex) natural property on which the moral property in question supervenes. This argument assumes that the subvening natural properties can be isolated so as to allow us to construct causal explanations solely in terms of such properties. However, if what was said here about the interconnectedness of the evaluative and the descriptive and about the supervenience principle is correct, then the aforementioned assumption is doubtful. If the descriptive and evaluative elements of some moral properties depends on the instantiation of other properties which we may be unable to represent in austerely nonevaluative terminology. But then one cannot just assume that evaluative features can be separated from nonevaluative ones and hold that the latter are doing all the causal work, whereas the former are causally irrelevant.

Critics of moral explanations can raise several objections to this line of argument. For instance, one can contend that thick ethical concepts do not in any way threaten the supervenience principle. More specifically, one can argue that even if the descriptive and evaluative elements of thick ethical concepts are not separable, this does not mean that when we focus on particular instances of thick concepts there will be no descriptive properties that make the use of a given concept appropriate. To illustrate this point, consider the concept of honesty. We have already seen in the previous section that the extension of this concept includes a wide range of actions with diverse descriptive properties. Arguably, any attempt to detach from this thick concept a purely descriptive concept that would cover all instances of honesty would be unsuccessful. Nonetheless, the objection would go, this does not mean that when we consider particular instances of honesty there are no descriptive properties that make the use of this concept appropriate. The relevance of these properties may be restricted to particular cases, but this would not change the fact that in each and every case we will be able to determine which descriptive properties make the use of the concept of honesty appropriate. I will later on provide a detailed answer to this objection. Yet, critics of moral explanations may simply claim that we should doubt that there is enough reason to reject the supervenience principle, since the supervenience of the moral on the natural is a conceptual truth. I will address this worry in what follows.

Recall from the previous section that the supervenience principle involves two claims: (a) the claim that, if an object possesses a moral property this can only be "because of" or "in virtue of" its possessing certain natural properties; and (b) the claim that no two objects can differ with respect to their moral properties without also differing with respect to their natural properties. I will hereafter refer to the dependence claim expressed by (a) as the *SD* principle, and to the covariance claim expressed by (b) as the *SC* principle.

The relationship between SD and SC is not always made explicit by the advocates of moral supervenience. Yet, most authors suggest that it is our commitment to SD that explains our commitment to SC.<sup>29</sup> Roughly put, the idea is the following. If objects owe their moral properties to their natural properties, i.e. if the instantiation of certain natural

CEU eTD Collection

<sup>&</sup>lt;sup>29</sup> See, for instance, Smith (2000: 108) and Zangwill (2006: 272–3)

properties is sufficient for the instantiation of certain moral properties, then it follows that no two objects can share all their natural properties without also sharing their moral properties. Other authors suggest that it is *SC* that should serve to elucidate SD.<sup>30</sup> In what follows, I will not assume that by rejecting one of the aforementioned principles one can automatically reject the other. I will be less concerned with the relationship between *SD* and *SC* than with showing that the powerful intuitions that seem to favor *SD* and *SC* can be preserved even if we reject these principles.

We seem to have a powerful intuition in favor of SD. One way to articulate this intuition would be to emphasize that, if something is good, this can only be because it is made good by other features. Suppose we say that a particular life is good. Then we must think that it is made good by other features it possesses, since nothing can be *just good*. Correspondingly, we believe that, if two objects differ with respect to their moral properties, there must be some further difference between them that would explain why this is so. (Note that this way of articulating our core pro-supervenience intuition makes our commitment to SC depend on our commitment to SD. However, as it will soon become clear, there is more than one way to articulate this intuition.)

The question is whether we should accept that, if an object possesses a certain evaluative property, this can only be in virtue of its possessing certain other *nonevaluative* properties. James Griffin believes that this question should be answered in the negative. He points out that the evaluative and descriptive elements of a concept such as "accomplishment" are intimately linked (1998: 25–9, 46). So, to claim that a particular life is good because it is a life of accomplishment means to claim that it is made good by this evaluative feature. If we ask what makes that particular life a life of accomplishment, we will have to refer to yet other evaluative features. Griffin emphasizes that the answer to this question would have to refer to the person's having done things that give life weight or point, and that a purely nonevaluative description of these things will be insufficient to make true the claim that they have the kind of importance needed to give life weight (2000: 300–1).

Michael Smith (2000) disagrees with Griffin. He sums up Griffin's position as follows: while Griffin admits that when we ascribe an evaluative property to an object we thereby incur an obligation to say what it is about the object in question that makes the ascription of that evaluative property appropriate, he claims that we can simply refer to some other, more specific, evaluative features. Smith (2000: 108–12) points out that the view according to which evaluative claims are made true by other, more specific, evaluative claims are made true by other, more specific, evaluative claims are made true by other, more specific, evaluative claims are "barely true". Yet, he finds it hard to make sense of this idea.

Griffin (2000: 299–300) asks why should we think that it is a conceptual truth that evaluative claims are always made true by other claims. He agrees that if something is good, it is because it is made good by *other* features. More specifically, he grants that to say that something is *just* good, would mean to violate the rules that govern the use of the word "good". However, in his view, the crucial question is whether saying that something is just good would violate not only the rules that govern the use of the word "good" in particular, but also the rules that govern the use of evaluative terms in general. Griffin argues that:

<sup>&</sup>lt;sup>30</sup> See Sturgeon (2006) and Strandberg (2008).

[A]nyone who uses the word 'good' in these contexts must be prepared to make claims about the interests that are the bases for the evaluation. But bases eventually come to an end. It is not a requirement of coherence that one can go on supplying bases for ever. Why might we not come to an end with a claim that is still evaluative? (Griffin, 2000: 300)

According to Griffin, the reason why Smith finds it difficult to accept that evaluative claims can be "barely true" has to do with the fact that such claims cannot be true independently of certain other nonevaluative claims. For example, a life cannot be good unless there is someone living the life. Griffin has no intention of denying this trivial point. However, as he rightly points out, Smith makes a stronger point. When Smith says that an evaluative claim is made true by other claims he means that these other (purely nonevaluative) claims are *sufficient* to make the evaluative claim true. However, insofar as Griffin is right that the evaluative and the descriptive are intimately connected, we have reason to reject the view that evaluative claims are always made true (in the sense specified above) by purely nonevaluative claims.

We have seen that one way to articulate our core pro-supervenience intuition is to say that, if something is good, this can only be because it is made good by other features. (Note that none of the considerations offered by Griffin offends against this formulation. However, this is only because this formulation refers to the evaluative term "good" instead of evaluative terms in general.) Another way to articulate this intuition would be to say that we need to have *reasons* for ascribing evaluative properties to things. Raz (1999) refers to this idea as the intelligibility of the domain of value. As he rightly points out, there is nothing *arbitrary* in this domain. This formulation allows us to explain why we can reject SD despite its initial plausibility. For the intelligibility of the domain of value implies that "there is an explanation for everything, an explanation for why what is good is good, what is bad is bad" (Raz, 1999: 220). While it is necessary that there is an explanation for everything, the explanations that we offer need not be couched in entirely non-evaluative terms. (Correspondingly, the concept of reason can be used to articulate the intuition that seems to favor SC. Instead of saying that there can be no evaluative difference without an underlying nonevaluative difference, one can say that when the same reasons apply, we must make the same moral judgment.)

Thus far, we have seen that the interconnectedness of the evaluative and the descriptive gives us reason to doubt the truth of SD. Yet, if we have reason to reject SD, it seems that we also have reason to reject SC. For if the moral does not necessarily depend on the natural, then it must be possible at least in principle to have two situations which are identical with respect to their natural properties, but which nevertheless call for different moral evaluations. This consequence of rejecting SD seems problematic, since one can argue that SC is an independently plausible principle. In the remainder of this section, I will attempt to show that rejecting SC is not as counterintuitive as it may first seem.

In order to prove that rejecting SC involves nothing counterintuitive, one may turn to Dancy's conception of "resultance". Dancy (2004: 85–9) distinguishes between supervenience and resultance. This distinction corresponds roughly to the distinction made here between a relation of covariance and a relation of (asymmetrical) dependence among moral and natural properties. However, Dancy emphasizes that whereas the supervenience base is supposed to include *all* the nonmoral features of an object, the resultance base includes only *some* of an object's nonmoral features. As he rightly points out, we do not think that all of an object's nonmoral properties contribute to that object's having a certain moral property; quite to the contrary. The resultance base for the wrongness of a given action does not include all of its nonmoral features, but only those features that make the action in question wrong.<sup>31</sup>

Dancy challenges the assumption that, if an action is made wrong by certain nonmoral features, then any action that possesses these features would also be wrong. He claims that, although two actions may share the features that make the first one wrong, these features may be prevented from making the second action wrong "because of variations that lie beyond the common resultance base" (2004: 87). He takes this point to be established by the distinction he draws between what he calls "favourers", i.e. nonmoral features that make actions have certain moral features, and what he calls "enablers", i.e. nonmoral features that merely enable other nonmoral features to make actions have certain moral features (2004: 38–41).

Dancy's considerations about favouring and enabling would provide an argument against a particular version of the supervenience principle. According to this version of the principle, it is necessary that, for each moral property M instantiated by an object, there is a set of natural properties N that the object has, and that M always accompanies N. Now, one can argue that it is this version of the supervenience principle that we should focus on. (In fact, this seems to be Dancy's suggestion. Dancy contends that if the supervenience base includes all the nonmoral features of an object, then the supervenience principle is uninteresting, even if true. In his view, the question that we need to answer is whether there is "any stable point on the way, any point, that is, short of the entire supervenience base, that would have the same result" (Dancy, 2004: 90). One suggestion would be to restrict the supervenience base to those nonmoral features that are part of the resultance base. Yet, the resultance base is too narrow to ensure that when the nonmoral features that make an object have a certain moral feature will recur they will have the same effect.<sup>32</sup>) However, this is not the version of the supervenience principle that I have been considering so far. As Dancy's discussion of supervenience shows, there is a difference between holding that (a) no two objects can share all their natural properties without also sharing their moral properties, and holding that (b) if an object has a moral property, any other object that shares the natural properties that make the first object have the moral property in question must also have that moral property. As defined above, the principle SC amounts to claim (b). And given that what many would find problematic about the rejection of SD is precisely that rejecting it would also lead to rejecting SC, what we need here is an argument against SC.

<sup>&</sup>lt;sup>31</sup> Dancy's view on supervenience bears close similarities to Griffin's view. Compare Dancy on the relation of "resultance" to Griffin (1998: 44–8; 2000) on the "relevance requirement". According to Griffin, without the relevance requirement, the supervenience principle is uninteresting, whereas with the relevance requirement it is doubtful. I take Dancy (2004) to express the same idea. Compare also Dancy (1996).

<sup>&</sup>lt;sup>32</sup> For analyzing the suggestion that the resultance base should be expanded so as to include both favourers and enablers, see Dancy (2004: 89–93). See also Strandberg (2008).

Once we accept the interconnectedness of the descriptive and the evaluative, the principle SC is thrown into doubt. Yet, we have seen above that there is a powerful intuition in favor of SC. In what follows, I will attempt to show that this intuition can be preserved even if we reject SC. More concretely, I will argue that there is a modified version of SC that does not in any way conflict with this intuition.

Recall that according to SC, no two objects can differ with respect to their moral properties without also differing with respect to their natural properties. We seem to have a powerful intuition in favor of this principle. One way to articulate this intuition would be to say that, if two objects differ with respect to their moral properties, there must be some further difference between them that would explain why this is so. The question is whether this further difference must always be specifiable in purely nonevaluative terms. Those who hold that there is no sharp distinction between the descriptive and the evaluative would most probably answer this question in the negative. More specifically, this answer would involve making the following three claims. First, while no two situations can call for different moral evaluations unless there is some further difference between them that would explain why this is so, it is possible that this further difference cannot be accounted for in purely nonevaluative terms. Second, SC needs to be amended to allow for the possibility just mentioned above. According to this modified version of the principle, no two situations can call for different moral evaluations unless there is some further difference (evaluative or nonevaluative) between them that would explain why this is so. Call this principle  $SC^*$ . Third, although it may seem that  $SC^*$  fails to do justice to our powerful pro-supervenience intuition, there are ways to articulate this intuition so as not to conflict with SC\*. As Dancy rightly points out:

[T]he thesis of the supervenience of the moral on the natural is held in place by the concept of a reason; it is effectively an expression of the thought that where exactly the same reasons are in place, one must make the same moral judgment. But since there now seems to be nothing to persuade us that the idea of a reason eventually takes us all the way down to the natural, since we will often more or less come to a halt at the thick, this claim about reasons does nothing to establish the supervenience claim. The latter now seems to me to be false. (Dancy, 1996: 279)

Once we accept that our core pro-supervenience intuition can be expressed, as Dancy suggests, in terms of reasons, it becomes doubtful that this intuition supports SC rather than a modified version of this principle such as the one proposed here.

Still, to deny that two situations can call for different moral evaluations only if they differ at the level of the natural is to make a bold claim, and one can argue that the burden rests upon those who defend such an eccentric position to show that we can conceive of two situations that are indistinguishable at the level of the natural without also being evaluatively indistinguishable. So the best way to proceed from this point on would be to provide an example that satisfies this description.

Consider the following example. Let W and W\* be two possible worlds. Suppose that in W, Bob is called to testify as a witness in a trial that takes place in a community different from his own. During his testimony, he is asked a question the answer to which would involve revealing certain aspects of his past that he is not proud of. Although the verdict would not be affected by whether he reveals these aspects of his life, the question happens to be formulated in such a way as to require giving such unpleasant details.

Suppose that Bob admits to having done things that both him and the audience find morally objectionable, and that it is his honesty that stops him from trying to avoid giving a straight answer to the attorney's question.

Now, let us suppose that Bob\* is Bob's identical twin in W\*. Suppose further that, in his testimony, Bob\* also admits to having done morally objectionable things. However, it is the fact that Bob\* wants others to believe that he is honest, and not his honesty, that explains why he chooses to reveal certain aspects of his past. (We can assume that, once the question referring to his past is brought up, Bob\* instantly realizes that the investigators possess information about his deeds, and that it is the fact that he wants to be perceived as an honest person that explains his testimony. By contrast, Bob's testimony has nothing to do with his realizing that the investigators are informed about his deeds.)

Undeniably, the two above-mentioned cases call for different moral evaluations. The question is whether we should admit that there is also a relevant *nonevaluative* difference between these cases. Let us start by noting that Bob's overt behavior is indistinguishable from Bob\*'s overt behavior. Yet, surely, one could reject behaviorism and still claim that it is not possible to get different moral evaluations without any underlying nonevaluative difference. So, the question that concerns us here is whether the difference between the two above-mentioned cases can be captured in purely nonevaluative terms, and not whether this difference can be accounted for in purely behavioral terms.

The difference between the two cases consists in the fact that, whereas Bob confesses that he has done morally objectionable things because he is honest, Bob\* confesses that he has done such things because he wants to be perceived as being honest. One can claim that there is no need to appeal to any evaluative terms in describing these cases. More specifically, one can argue that the term "honest" can be replaced by other, nonevaluative terms without any explanatory loss. For instance, instead of saying that Bob\* acted as he did because he wanted to be perceived as being honest, we could say that he wanted to enjoy some of the advantages resulting from his acting in a certain way. advantages that we would have no trouble describing in nonevaluative terms. However, it is not at all clear what these advantages could be. While it is true that being perceived as an honest person has certain advantages (such as bettering one's prospects of participating in future cooperative enterprises), given that we have assumed from the beginning that the trial at which Bob\* is summoned as a witness takes place in a community different from his own, it is doubtful that a decision to give the impression of an honest person is to be explained by his thinking about such advantages. Still, one can point out that there is one obvious advantage that would result from Bob\*'s giving the impression of an honest person. Since his testimony is given under oath, the consequences of not telling the truth would be quite serious. Although Bob\*'s decision to act as he did can be understood as a decision to avoid the consequences of perjury, this example is formulated in such a way as to leave other possibilities open. (The assumption is that whether Bob\* reveals unpleasant details about his past does not affect the verdict, and that the attorney is not so intent on exposing such details. So Bob\* could just as well answer the question without specifically mentioning these details.) There seems to be nothing unreasonable about the assumption that Bob\* acts as he does not because he is thinking about the advantages that would result from his action, but simply because he wants to be perceived as an honest person. At this point, one can argue that, if Bob\*'s confession can be explained solely by reference to his wanting to be perceived by others as an honest person, then this must be because he wants others to like him. Yet, it is doubtful that the evaluative term "honest" can be eliminated by referring to Bob\*'s wanting to be liked by other people. To see why this is so, think about the fact that by confessing that he has done things that people in the audience disapprove of, Bob\* exposes himself to the risk of not being liked by them. So, saving that Bob\* acts as he does simply because he wants other people to like him would fail to capture a distinctively evaluative element (i.e. his preferring to be disliked for having done a certain kind of morally objectionable things rather than for being dishonest).<sup>33</sup> The upshot of all these considerations is that we cannot give an adequate account of why Bob\* acts as he does in purely nonevaluative terms. Now, going back to Bob, we have reason to doubt that attempts to substitute some nonevaluative terms for the evaluative term "honest" can be successful. At first glance one might think that there is a somewhat narrow, nonevaluative description of honesty that would serve the limited purpose of this example. Thus, one might refer, for instance, to Bob's willingly telling the truth instead of his being honest. However, once we consider Bob\*'s case, it becomes obvious that there is more to Bob's being honest than this nonevaluative description can capture, given that Bob\* also willingly tells the truth.

What does this example tell us? The two cases discussed here call for different moral evaluations. The question that we have been asking is whether we should admit that there is also a relevant *nonevaluative* difference between these cases. We have seen that the only difference between these two cases consists in the fact that, whereas Bob confesses that he has done morally objectionable things out of honesty, Bob\* confesses that he has done such things because he wants to be perceived as an honest person. Furthermore, we have seen that we cannot account for this difference in purely nonevaluative terms.

There is, then, a counterexample to principle *SC*, since the cases considered above are indistinguishable at the level of the natural without also being indistinguishable evaluatively. Yet, appearances notwithstanding, rejecting *SC* does not offend against our core pro-supervenience intuition. If Dancy is right that the best way to articulate this intuition is to say that where the same reasons apply, we must make the same moral judgment, then the cases considered here do nothing but support this intuition. For there is a relevant difference between these cases (even though this difference cannot be captured in purely nonevaluative terms), and this is why they call for different moral evaluations. Had they been similar in all respects, surely they would have called for the same moral evaluation.

<sup>&</sup>lt;sup>33</sup> It can be objected that if one prefers to be disliked for having done a certain kind of morally objectionable things rather than for being dishonest, then it certainly looks like the person in question is an honest person. This claim would undermine the present argument by showing that basically there is no difference (evaluative or nonevaluative) between the two cases considered here. However, the formulation above does not entail that Bob\* acts out of honesty. One possibility is that Bob\* prefers to be disliked for having committed moral mistakes in his past rather than for committing such mistakes at the present time. Another possibility is that, although he prefers to be known for having committed certain kinds of moral mistakes rather than for being dishonest, he would at the same time prefer that these mistakes were unknown to anyone.

Presumably, those who defend the distinction between the descriptive and the evaluative would argue that we can account for the difference between the two cases discussed above in purely nonevaluative terms. My aim was to cast doubt on the idea that we can substitute some nonevaluative terms for the evaluative term "honest". There is, however, a further point that I would like to make here. It seems that whether or not one agrees that we have reason to abandon the principle *SC* actually comes down to whether or not one believes that it is always possible to adequately represent any aspect of a situation by using purely nonevaluative terms.

This is a contentious point. Nonetheless, my point is that rejecting the principle that the moral necessarily depends on the natural does not lead to *further* counterintuitive consequences. So, one cannot object to rejecting the principle *SD* on the grounds that this would have as an unwelcome consequence the rejection of *SC*.

Now, let me briefly go back to an objection raised at the beginning of this section. According to this objection, even if the evaluative and descriptive elements of thick ethical concepts are intimately connected, this does not mean that when we focus on particular instances of thick concepts we will be unable to determine which descriptive properties make the use of a certain concept appropriate. To put it differently, one can hold that (a) the descriptive properties that make the use of a certain concept appropriate are so diverse that it is only by lumping them under an evaluative concept that their unifying characteristics become evident, and that (b) this is all there is to the interconnectedness of the evaluative and the descriptive. Yet, surely the proponents of the view that the evaluative and descriptive elements of thick ethical concepts are indissolubly linked want to make a stronger claim. For instance, Dancy writes that: "The 'evaluation' partially determines the 'description', since the only way to specify the nature of the object of which we approve was to say that it has these or those features in the (or a) right way" (1996: 276). In other words, the approval is not bestowed for having certain descriptive properties, but for having those properties in the right way. I hope to have provided some content to the idea that in order for an object to come under a thick concept, having certain descriptive properties is not enough. The example discussed above shows that what makes the use of the concept of honesty appropriate is not the mere presence of certain descriptive properties, but rather their being present in the right wav.

At this point, it is worth considering a recent attempt to defend the distinction between the descriptive and evaluative components of thick ethical concepts. Andrew Payne (2005) argues that this distinction can be preserved by introducing the notion of "thick description". Thick descriptions convey information about the agent's intentions, beliefs and desires. The notion of thick description would allow us to capture what different instances of a thick concept have in common without simultaneously crossing over to evaluation. On this view, it is in virtue of their intentions, beliefs and desires that agents and their actions would fall under a thick concept. Now, the question is whether Payne is right in thinking that each thick ethical concept has a thick descriptive component and that satisfying such a thick description is all that is needed for an agent or an action to fall under a thick concept. The example considered above suggests that the answer to this question is negative. Certainly, the view that thick concepts have a descriptive component that is specifiable in terms of the agent's intentions and desires is more promising than the view that this descriptive component is specifiable in purely behavioral terms. However, there does not seem to be any thick descriptive component of honesty that would be satisfied by Bob's case, but not by Bob\*'s case. As we have seen, we cannot account for the difference between these two cases in purely nonevaluative terms. Consequently, we should conclude, with Dancy, that the descriptive properties that make the use of a concept like honesty appropriate must be present in the right way.

Those who are suspicious of the idea that there is no thick descriptive component of honesty that would be satisfied by Bob's case, but not by Bob\*'s case, should consider some further suggestions about how to account for the difference between these cases in nonevaluative terms. Perhaps instead of saying that Bob\* wanted to be perceived as an honest person we could say that, unlike Bob, Bob\* had the intention to deceive the audience. However, this suggestion does not seem to be helpful given that, had Bob\* decided to avoid giving a straight answer to the attorney's question, this decision could also have been explained by his intending to deceive the audience. So, saying simply that Bob\* had the intention to deceive the audience does not offer an adequate representation of why he acted as he did. At this point, one might raise the following objection. What matters is not that we cannot explain why Bob\* decided to make a confession rather than not to make it by reference to his intending to deceive the audience. What is important is that his intention to deceive allows us to account for the difference between his case and Bob's case in purely *nonevaluative* terms. We can say that, *whatever else he might have* wanted, Bob\* intended to deceive the audience. This provides a clear contrast to Bob's case.

In response to this objection, let me first emphasize that it is doubtful that the concept of honesty has a thick descriptive component that would refer to telling the truth without any intention to deceive. According to Payne, it is necessary that all instances of honesty satisfy the relevant thick description. However, it seems that there are instances of honesty which fail to satisfy the above-mentioned thick description. Recall that Railton emphasized that honesty requires different things from a person who is describing an illness to a child, or from a person who has promised confidentiality to someone else, or from a person who plays a friendly poker game. Surely, one can argue that "deceiving" can mean different things and that only some of these things can be associated with an honest behavior. Nonetheless, it is doubtful that we can make such distinctions without appeal to any evaluative terms. Now, even if there is no thick description of honesty that would reveal what is common to all instances of honesty, the objection that it is possible to account for the difference between the cases of Bob and Bob\* in purely nonevaluative terms still stands. One can say that the difference between the two cases comes down to the fact that, unlike Bob, Bob\* intended to deceive the audience. Yet, a modified version of this example would allow us to answer to this objection. Suppose that Bob chose to omit from his confession any details that would present in a bad light a certain person that he met during his youth. If so, we could say that Bob intended to mislead the audience into thinking that a person whose name was mentioned many times during the trial was a morally admirable person. Certainly, one can point out that there are different understandings of "deceiving" and that the sense in which Bob intends to deceive the audience is not the same as the sense in which Bob\* intends to deceive the audience. Maybe Bob has a noble aim, such as protecting the image of someone who was not present to defend herself. But then again, it is doubtful that we can distinguish between different understandings of deceiving without appealing to any evaluative terminology.

Another suggestion would be to say that, by contrast to Bob's decision to make a confession about his past, Bob\*'s decision to make such a confession is the result of a cool calculation. Yet, we have seen that we can complicate our example so as to make it plausible that Bob's decision is also based on a calculus of consequences (i.e. we can suppose that Bob decided to make a confession only after reflecting upon whether his doing so would present other persons in a bad light). Then perhaps we can say that Bob\*'s decision to make a confession is the result of a cool self-interested calculation. However, as we have already seen, it is not at all clear how making such a confession would serve Bob\*'s interest. If, on the other hand, we take into consideration only his desire to be liked by others, then we are back with the dilemma presented above (i.e. it is not at all clear why Bob\* would prefer to be disliked for having committed a certain kind of moral mistakes rather than for being dishonest).

Let me take stock. In this section, I have argued that the realist can win the debate about whether moral properties figure ineliminably in causal explanations by expressing doubts about the truth of the supervenience principle. I have attempted to prove that the intimate manner in which the evaluative and the descriptive are interconnected gives us reason to doubt the truth of this principle.<sup>34</sup>

My aim was to show that proponents of both naturalistic and nonnaturalistic versions of moral realism can endorse the claim that moral properties can figure in causal explanations. Reductive naturalism would have no trouble accommodating this claim, given that, on this view, moral properties are either identical to (or else reducible) to natural properties. The more interesting question is whether non-reductive naturalism or nonnaturalism can also accommodate this claim. My main concern was to establish that ethical nonnaturalists can endorse the claim that moral properties play genuine causal explanatory roles. Nonnaturalists can argue that, given the interconnectedness of the evaluative and the descriptive, there is no principled reason for claiming that the causal explanatory work is always done by the subvening nonevaluative properties. It is worth emphasizing that most nonnaturalists agree with the antirealist claim that moral

<sup>&</sup>lt;sup>34</sup> According to one prominent version of the supervenience argument against moral realism, a weaker claim about supervenience is true (i.e. the claim that, for each moral property M instantiated by an object, there is a natural property N that the object has, and that M always accompanies N as a matter of metaphysical necessity), whereas another, stronger claim is false (i.e. the claim that M always accompanies N as a matter of *analytic necessity*). Allegedly, the realist - as opposed to the antirealist - cannot explain why the weaker supervenience claim holds true, given that the stronger one does not. (See Blackburn (1971; 1985b). For useful discussion, see also Dreier (1992)). There are several possible responses to this argument, depending on which version of realism one accepts. According to reductive naturalism, this argument is question-begging. This is because reductive naturalists believe that the stronger supervenience claim is in fact true. Non-reductive naturalists can point out that their theory has no trouble explaining why the weaker supervenience claim holds true (given that moral properties are natural properties), while the stronger one does not (given that, on this view, there are no analytic relations between moral and nonmoral truths). What about moral nonnaturalism? According to Blackburn, what makes it hard for the realist to answer to the supervenience argument is his committed to the lack of entailment thesis, according to which no set of nonmoral truths entails any particular moral truth. Yet, some nonnaturalists hold that realists should give up the lack of entailment thesis. (See Shafer-Landau (2003: 84-98)). The suggestion made here is that (nonnaturalist) realists can respond to this argument by raising doubts about whether the relevant supervenience claims can be expressed in a way that is both interesting and neutral among different metaethical doctrines.

properties are causally inefficacious. Presumably, this is because they believe that to hold otherwise would mean to accept that moral properties are natural properties.<sup>35</sup> However, if the argument developed here is correct, then holding that moral properties are causally efficacious need not imply that they are natural properties. The following excerpt from Cuneo offers a suggestion about how to extend the present argument:

A person's virtue consists in the capacities to notice what is morally salient, to interpret those signs *correctly*, and to be motivated *appropriately*. A person's virtue, then, is a richly evaluative thing that consists in being reliably disposed to respond to reasons of various kinds appropriately. (Cuneo, 2007b: 357)

Cuneo advocates the view that moral virtues are the most likely candidates among moral facts for playing causal explanatory roles. Insofar as he is right in claiming that a person's virtue, or its manifestations, cannot be reduced to a set of purely descriptive properties, there seems to be no incompatibility between embracing nonnaturalism and holding that moral traits of character are causally efficacious.

I will not deal here with the question whether any of the considerations presented in this section would offend against non-reductive naturalism. However, it is worth noticing that Sturgeon himself advocates a line of defense against the charge that moral facts are causally inert similar to the one proposed here. Whereas considerations about moral supervenience were central to his initial reply to Harman, in his latter writings Sturgeon points out that the supervenience principle is more doubtful than it is commonly supposed to be.<sup>36</sup> In his view, a reason for doubting the truth of the supervenience principle is that there may be natural properties which make an evaluative difference, but for which we lack nonevaluative terminology. So, Sturgeon suggests that, in response to the claim that all the explanatory causal work is actually done by the natural facts on which moral facts supervene, the realist should express doubts about whether it is possible to come up with a version of the supervenience principle that is both philosophical interesting and neutral between different metaethical positions.

## 3.5 Vindicatory explanations

So far, I have argued that in responding to the charge of explanatory impotence both naturalists and nonnaturalists can claim that some moral facts figure ineliminably in causal explanations. The suggestion made here is that moral traits of character are the most plausible candidates for being causally efficacious. We have already seen that according to Sturgeon, one of most ardent naturalist proponents of the causalist position, facts about one's moral character can plausibly be cited as part of a reasonable explanation of one's deeds. Yet, it is important to notice that there are also nonnaturalists who embrace the above-mentioned suggestion. For instance, David Wiggins (1990–91; 1996) holds that moral properties are nonnatural properties that need not pull their weight

**CEU eTD Collection** 

 <sup>&</sup>lt;sup>35</sup> For this point, see Sturgeon (2006: 256, n6).
<sup>36</sup> Compare Sturgeon (1985) to Sturgeon (2006: 247–9).

in a scientific account of the world. However, Wiggins points out that some moral facts actually do figure in causal explanations of nonmoral facts. He stresses that:

The courage or negligence or callousness or considerateness of a person x can make a vast difference to outcomes other than the beliefs and motives of persons. It can make a causal difference to the fate of persons either identical with x or different from x. And that which affects thought which modifies outcomes does affect outcomes. (Wiggins, 1996: 48)<sup>37</sup>

So, according to Wiggins, facts about one's moral character can figure in causal explanations. However, Wiggins's position is atypical for someone who endorses a nonnaturalistic version of moral realism. The argument in the previous section was intended to show that, despite what is commonly assumed by both naturalists and nonnaturalists, claiming that some moral facts can play a genuine causal explanatory role is not at odds with endorsing nonnaturalism.

In this section, I will consider Wiggins's response to the argument from explanatory impotence. As I have already emphasized, most nonnaturalists agree with the antirealist claim that moral facts are causally inert, while arguing that realism can accommodate this fact.<sup>38</sup> What sets apart Wiggins's position is that he claims that some moral facts are causally efficacious. Yet, there are also important differences that set apart Wiggins's position from the position adopted by Cornell realists. Most importantly, Wiggins thinks that the *Causal Criterion* of justified belief should be rejected in favor of the weaker *Explanatory Criterion*. In addition, he holds that moral facts do not cause our beliefs about them, although they can cause other nonmoral facts.

Wiggins rejects the causal theory of knowledge and the corresponding *Causal Criterion* of justified belief and holds that moral facts figure in a special class of explanations, which he calls "vindicatory explanations". Yet, he does not go to great length to justify the rejection of the *Causal Criterion*. Instead, he focuses on showing how vindicatory explanatory of our moral beliefs work. However, enough has been said so far to throw doubt on the idea that the *Causal Criterion* of justified belief is more appropriate than the *Explanatory Criterion*. As already emphasized, we have reason to believe that the *Causal Criterion* is too strong. As Sayre-McCord rightly points out that:

[E]mpirical generalizations (like "all emeralds are green"), as well as natural laws (like the first law of thermodynamics), would fall victim to the Causal Criterion. Although these generalizations may help *explain* why we experience what we do as we do, they *cause* none of our experiences. That all emeralds are green does not cause a particular emerald to be green, nor does it cause us to see emeralds as green. (Sayre-McCord 1988b: 266, emphases in original)

Thus, the *Causal Criterion* would discard too many of our beliefs. Furthermore, recall from section 3.3 that causal and explanatory reasoning are closely intertwined. We have

<sup>&</sup>lt;sup>37</sup> Wiggins's argument is directed against those who hold that all attempts made by realists to prove that moral facts play an explanatory role fall prey to the same objection, namely that the explanatory role attributed to moral facts seems to be especially tailored for a particular class of phenomena (i.e. moral beliefs). For this line of reasoning, see Wright (1992: ch. 5) and Leiter (2001: 82–3).

<sup>&</sup>lt;sup>38</sup> See, for instance, Nagel (1986), Thomson (1996), McGinn (1997: 35–60), and Audi (1997). For a similar view, see also Dworkin (1996*b*).

already seen that, according to one of the most sophisticated analyses of our inductive practices, explanatory considerations are an important guide to causal inference, and that this helps to allay at least some of the worries that motivate the rejection of the *Explanatory Criterion* (e.g. the worry that explanatory reasoning cannot reliably take us towards the truth).<sup>39</sup> Another way of proving that the worries that motivate the rejection of the *Explanatory Criterion* are ill founded would be to offer some positive argument to the effect that there can be a reliable belief formation process even in the absence of a causal chain connecting the knower and the object known. This is precisely the strategy pursued by Wiggins.

At this point, one might ask a legitimate question: if it turns out that we should reject the Causal Criterion in favor of the Explanatory Criterion, why should the realist insist that moral facts can figure in causal explanations? Recall that the reason why Sturgeon was interested in proving that facts about one's moral character form part of a reasonable explanation of one's actions was that he thought he has found a way of proving that moral facts figure in the causal explanation of our moral beliefs. On this view, the fact that an individual possesses a certain trait of character, such as cowardice or compassion, causes him to act in a certain way. In turn, his actions cause us to believe that he is coward or compassionate. Wiggins shares Sturgeon's belief that some moral facts figure in the causal explanation of nonmoral facts, while denying that moral facts causally explain our beliefs about them. Presumably, the reason why Wiggins still thinks it is important to emphasize that some moral facts are causally efficacious has to do with the fact that, in his view, the hidden appeal of the argument from explanatory impotence derives from the idea that moral facts are intrinsically queer (Wiggins, 1990-91: 80). An observation made by Colin McGinn may be especially relevant at this point. According to him, when having to deal with the argument from queerness the most important question to address is "whether there is any pretheoretical sense in which moral properties are intrinsically queer" (McGinn, 1997: 19, emphasis in original).<sup>40</sup> Undoubtedly, some will think that facts which are causally inert (or, to put it differently, facts which do not contribute to the explanation of any observable phenomena) are queer, whereas others will insist that there is nothing queer about morality, mathematics, logic or modality. And, as in the case of many other philosophically interesting concepts, the prospects of coming up with a neutral definition are not that bright. Most probably, arguments to the effect that the *Causal Criterion* is too strong and that we should opt instead for the

<sup>&</sup>lt;sup>39</sup> As we will see bellow, the idea that facts which are causally inert are intrinsically queer can also serve as a basis for rejecting the *Explanatory Criterion* in favor of the *Causal Criterion*.

<sup>&</sup>lt;sup>40</sup> One can argue that claiming that facts which are causally inert are queer comes down to claiming that we are not justified in believing in facts which are causally inert. There is, however, a significant difference between the two aforementioned claims. According to the latter claim, we cannot have epistemic access to things that are not part of any causal chain. So, if moral facts are outside any causal chain, they are epistemically inaccessible. In order to refute this claim, one needs to prove that there is a reliable belief formation process even in the absence of a causal interaction with moral facts. Now, if one were to prove that *some* moral facts that are causally efficacious, this would do nothing to establish that we are justified in believing in those moral facts that are causally efficacious. By contrast, in order to refute the claim that moral facts are made queer by their causal inefficacy, it is enough to prove that *some* facts are causally efficacious. If, on the other hand, one would come up with an argument to the effect that even if we bear no causal connection to moral facts look less queer in the eyes of antirealists.
*Explanatory Criterion* would fail to convince those who are already predisposed towards naturalism. If this is so, then even if the realist thinks that the *Explanatory Criterion* of justified belief is more appropriate, he has reason to consider the question whether moral facts can figure in causal explanations, and if it turns out that they actually do, this would strengthen his reply to the argument from explanatory impotence.

Let me say again that the reason why I am interested in the argument from explanatory impotence has to do with the fact that, as I have shown in the previous chapter, whether the realist can successfully reply to the argument from disagreement depends in part on whether he can successfully reply to the argument from explanatory impotence. Hence, what matters for my present purposes is whether (convergence of) moral belief can suitably be regarded as the result of a reliable epistemic mechanism. As already mentioned, ethical nonnaturalists generally take this question to be distinct from the question whether moral facts possess any causal efficacy. Although I agree with this position, I have attempted to show that some moral facts are causally efficacious, and that admitting this fact does not commit us to ethical naturalism. My main reason for proceeding this way has to do with the fact that, apparently, those who are already inclined towards naturalism are reluctant to accept arguments aimed at showing that the Explanatory Criterion is more appropriate than the Causal Criterion, and that moral facts satisfy the constraints of the former criterion. This reluctance may be explained by a tacit adherence to the idea that facts which possess no causal efficacy are intrinsically queer. But there may be also other explanations for this reluctance. For instance, faced with an argument to the effect that, even though moral facts are causally inert, they can explain (non-causally) our beliefs about them, one may be unable to escape the feeling that the explanatory role that moral facts are supposed to play seems to be especially tailored for a particular class of phenomena (i.e. our moral beliefs).<sup>41</sup> So, even if arguments to effect that some moral facts are causally efficacious have no direct bearing on the question whether we should accept the *Explanatory Criterion*, or whether moral facts actually meet this criterion, such arguments can help allay suspicions like the ones referred to above. The argument from the previous section is intended to convince those who entertain such suspicions.

Now, let me turn to Wiggins's view of vindicatory explanations. Wiggins (1987b; 1990–91; 1996) challenges the assumption that causal interaction is a necessary condition for knowledge. He argues that we can have epistemic access to moral facts although we bear no causal connection to them. Wiggins (1990–91: 66) points out that to claim that one comes to *know* that p is to claim that it is not merely accidental that one comes to

<sup>&</sup>lt;sup>41</sup> For raising the problem of the explanatory narrowness of moral facts, see Wright (1992: 191–201) and Leiter (2001: 82–3). Wright explicitly states that he would have nothing against dropping the causal criterion, as long as it can be proved that the facts that we want to construe realistically play a wider explanatory role. (He writes that: "[...] it is not my intention that the Wide Cosmological Role constraint should be satisfiable only by *causally active* states of affairs, nor even that the explanations involved have to be causal. The overarching point, remember, is that there be a *wider range of intelligible and legitimate uses of the relevant state denoters* [...]". (1992: 198, emphases in original)). In his view, Harman was mistaken in thinking that the relevant explananda for determining whether moral facts can be construed realistically are moral beliefs. Wright argues that the crucial question is not whether moral facts "feature in the *best* explanation of our beliefs about them, but of *what else* there is, other than our beliefs" (1992: 196–7, emphases in original).

believe that p. In other words, one comes to know that p if and only if one comes to believe that p precisely because p. This leads to the claim that the fact that p should figure in the best explanation of one's coming to believe that p. So much is uncontroversial. According to the causal theory of knowledge, the only sense we can make of the claim that "one comes to believe that p precisely because p" is that there is some sort of a causal chain relating one to the fact that p, so that the fact that p is the causal source of one's belief that p. However, Wiggins argues that, while it is necessary that the fact that pshould play a role in explaining one's coming to believe that p if the belief in question is to count as knowledge, the explanation involved does not have to be causal. He claims that some of our beliefs admit of what he calls "vindicatory explanations". This sort of explanations conforms to the following schema:

[F]or this, that and the other reason (here the explainer specifies these), there is really nothing else to think but that p; so it is a fact that p; so given the circumstances and given the subject's cognitive capacities and opportunities and given his access to what leaves nothing else to think but that p, no wonder he believes that p. (Wiggins, 1990–91: 66)

Wiggins illustrates how vindicatory explanations work by means of the following example:

My son (aged nine) believes, and all other boys and girls in his class at school believe, that 7 + 5 = 12. The best explanation of why they all believe this is not that have learnt and taken on trust the one truth '7 + 5 = 12' but (I hope and believe this): (i) As can be shown by use of the calculating rules (and could in the end be rigorously demonstrated), it is a fact that 7 + 5 = 12. There is nothing else to think but that 7 + 5 = 12.

(ii) The best explanation of my son and his classmates' shared belief is that they are going by the calculating rule that shows there is nothing else to think but that 7 + 5 = 12. If there is nothing else to think, then no wonder that, if their beliefs are answerable to the calculating rules, they agree in the belief that 7 + 5 = 12. (Wiggins, 1990–91: 67–8)

In his discussion of the argument from explanatory impotence, Crispin Wright (1992: 185) contributes a useful clarification of Wiggins's idea of vindicatory explanations. Wright points out that maintaining that a subject's belief that p admits of a vindicatory explanation involves two claims: (i) the claim that, when properly applied, the procedures of assessment in the relevant area of discourse leave no other option but the verdict that p; and (ii) the claim that the subject's belief that p is formed and guided by the application of the relevant procedures. Wright goes on to emphasize that even though vindicatory explanations cite the causes of the subjects' holding certain beliefs, these causes do not consist in the truth-conferring states of affairs. He writes that:

[W]hat causes the children's belief that 7 + 5 = 12 is what happens when they appropriately attentively apply the relevant arithmetical rules. What is cited in the vindicatory explanation is not the fact that 7 + 5 = 12, but the fact that, in the light of proper application of relevant rules, there is nothing else to think but that 7 + 5 = 12 [...]. (Wright, 1992: 187)

#### As Wiggins himself emphasizes:

Vindicatory explanations are causal explanations but the causality that they invoke is not one that holds between minds and values or between minds and integers. That would be a gross misunderstanding of what is got across by the explanatory schema exemplified by: 'There's nothing else to think but that 7 + 5 = 12. So no wonder they think that 7 + 5 = 12'. (Wiggins, 1990–91: 80)

According to Wiggins, moral beliefs admit of vindicatory explanations. To stick to an example used by him, why is it that we think that slavery is unjust and insupportable? Arguably, the belief that "slavery is unjust and insupportable" is formed and guided by the procedures of assessment appropriate to moral discourse, procedures which, when properly applied, leave *nothing else to think* but that slavery is unjust and insupportable. In this case (as opposed to other more controversial cases), it seems that the price of thinking anything else is to have opted out altogether from "the point of view that conditions our understanding of 'unjust' or 'insupportable'" (Wiggins, 1990–91: 71). Wiggins points out that what is essential to the sort of realism he espouses is that there will be a significant number of moral judgments that would command convergence in belief, and that this convergence admits of vindicatory explanation.

Before moving on to the next chapter, let me highlight the advantages of embracing a position like the one defended by Wiggins.

(1) We have seen that the argument from explanatory impotence relies on the assumption that, unless moral facts exert a *causal* influence on our perceptual apparatus, there is no plausible story to be told about how we can have epistemic access to them.<sup>42</sup> Wiggins's account of vindicatory explanations shows that this assumption is mistaken.

Wiggins argues that at least some moral and mathematical beliefs admit of vindicatory explanations. Yet, in defending the idea that there is nothing queer about moral epistemology, he does not only rely on the analogy between moral and mathematical beliefs. He also attempts to show that "there need not be any dichotomy between getting it right in matters of ethics, however distinctive that is, and getting it right in matters of fact" (2005: 7, emphasis in original). As Wiggins points out, the charge that realists need to postulate a strange faculty of intuition has to do with the supposed remoteness from perception of nonnatural properties. Mackie thought that believing in the existence of values commits one to the existence of nonnatural properties. The most common response to his argument from queerness is that he failed to consider the possibility that moral properties are natural properties. However, this is not the only strategy available to the realist in responding to this argument. As Wiggins emphasizes, the realist can claim that moral properties are nonnatural properties, while at the same time insisting that nonnatural properties are less strange than Mackie seemed to suppose. According to Wiggins (1996: 48, 2005: 9), a nonnatural property is simply property that need not pull its weight in a scientific account of the world. Now, Wiggins claims that the idea that our grasp of natural properties is better than our grasp of nonnatural properties is rather a myth. His argument against the claim that there is

<sup>&</sup>lt;sup>42</sup> See Williams (1985*a*: 136). For a suggestion to the same effect, see also Harman (1986: 62).

something queer about the epistemology of nonnatural properties that are value properties draws on Putnam's criticism of the fact-value dichotomy. Putnam asks:

How could there be 'value facts'? After all, we have no sense organ for detecting them? [...] Consider the parallel question: 'How could we come to tell that people are *elated*? After all, we have no sense organ for detecting elation' [...] Once I have acquired the concept of elation, I can see that someone is elated [...] Perception is not innocent; it is an exercise of our concepts. (Putnam, 2003: 102–3)

And Wiggins elaborates:

This is to say that once you have the concept of elation, you know what to look for. In looking for that, you can use any kind of perception or any mode of investigation that suits the case. Similarly then consider the predicate 'considerate'. That which marks out or delimits or descries or discriminates the property of considerateness in acts or attitudes or human characters is an essentially ethical interest, in the pursuit of which we can deploy any kind of perception or any mode of investigation or any associated concept that suits the case. (Wiggins, 2005: 10)

So, Wiggins does not only offer an account of *how* we can have epistemic access to moral facts, he also helps put to rest the worry that the process by which come to know such facts is different from our ordinary ways of knowing everything else.

(2) This brings out a further advantage of adopting a position like the one Wiggins defends. Wiggins's emphasis on the idea that the dichotomy between getting it right in matters of ethics and getting it right in matters of fact is not tenable helps to forestall an objection that envisages the nonnaturalists' most common responses to the charge of explanatory impotence.

Noncognitivists object to the idea that naturalistic constraints on epistemology and ontology which are otherwise accepted should be dropped when it comes to the subject matter of morality.<sup>43</sup> Consequently, they are unwilling to accept responses to the argument from explanatory impotence which are built around the claim that there is something distinctive about morality which explains why commonly accepted naturalistic constraints fail to apply to it. However, reaching this dead end can be avoided by shifting the emphasis from what is distinctive about moral judgments to what is common to moral and factual judgments. This is precisely the strategy that Wiggins pursues when he argues that there need not be any dichotomy between "getting it right in matters of ethics, *however distinctive that is*, and getting it right in matters of fact" (2005: 7, emphasis added).

(3) According to Wiggins, although moral facts do not cause our beliefs about them, they can cause other (nonmoral) facts.<sup>44</sup> We have already seen that the causal efficacy of moral

<sup>&</sup>lt;sup>43</sup> See Leiter (2001: 89–90). For an argument along the same lines, see also Svavarsdóttir (2001: 182–5).

<sup>&</sup>lt;sup>44</sup> This view is shared by Zangwill (2006). It is also worthwhile noticing that Wiggins's view on the process of belief formation bears similarities to Wright's view on the subject. Wright casts doubt on the idea that scientific beliefs are causally explained by the states of affairs that purportedly confer truth upon them. He holds that "the best explanation of a theorist's holding particular scientific theoretical beliefs should

facts strengthens the case for moral realism. The mere fact that moral facts can figure in causal explanations helps to dispel the idea that moral facts are intrinsically queer, whereas the fact that moral facts explain things other than our beliefs about them helps to allay the suspicion that the explanatory role they are supposed to play is especially tailored for a particular class of phenomena.

(4) Some might think that by defending the claim that *some* moral facts can figure in causal explanations, one cannot succeed in dispelling the force of the argument from explanatory impotence. Yet, we have already seen that coming up with an argument to the effect that *some* moral facts are causally efficacious is all that is needed to allay suspicions like the one referred to at (3). Now, it is worthwhile noticing that claiming that *only* some moral facts are causally efficacious might in the end turn to the realist's advantage.

A strategy commonly deployed by those who claim that moral facts possess no causal powers is to argue that moral facts of a certain kind cannot be causally efficacious and then to conclude that moral facts, as such, are causally inert. However, as Cuneo (2007b: 355) rightly points out, this strategy is questionable. Those who adopt it do not consider the question whether the realist must be committed to more than just the thesis that *some* moral facts are causally efficacious, or the question whether the kinds of moral facts that are claimed to be causally inert are the most plausible candidates for being causally efficacious. Cuneo goes on to argue that the view according to which moral facts possess causal powers is put in the best light by claiming that moral virtues and vices are causally efficacious. (Wiggins agrees with Cuneo on this point.) Once it becomes clear that the realist need not be committed to more than just the claim that *some* moral facts are causally efficacious (for instance, because in answering to the charge of explanatory impotence, it would be enough to show that the above-mentioned claim is true), the realist does not have to account for the fact that certain kinds of moral facts are seemingly causally inert. Examples involving causally inert moral facts simply lose their relevance.

confine consideration to his scientific inheritance, his observations, certain psychological laws, and features of his background psychology" (1992: 190). See also *supra* note 4.

# Chapter 4 Value incommensurability

#### 4.1 Value incommensurability

The notion of value incommensurability is usually contrasted with the notion of conceptual incommensurability, which has its origins in the writings of Thomas Kuhn.<sup>1</sup> The latter notion is intended to capture the difficulties of evaluation across conceptual schemes, ways of life or ethical outlooks embedded in different cultures, difficulties that arise due to a breakdown in translatability or to the lack of shared standards of rationality. By contrast, the notion of value incommensurability points to the difficulties that need to be dealt with when engaging in a process of evaluation within a given conceptual scheme, way of life or ethical outlook.<sup>2</sup>

In what the notion of value incommensurability is concerned, there are two main ideas in play.<sup>3</sup> First, "incommensurability" refers to the lack of a single scale of units of value on which two given items can be precisely measured. Ruth Chang casts doubt on the significance of this notion of incommensurability.<sup>4</sup> She rightly points out that comparisons do not require precise measurement in terms of common units of value. Where cardinal ranking fails, ordinal ranking may still be possible. We can judge, for instance, that *A* is a better mathematician than *B* without judging that *A* is twice as good a mathematician as *B* is.

Second, "incommensurability" refers to the idea that certain items cannot be ranked against each other on an ordinal scale. In Joseph Raz's words:

A and B are incommensurate if it is neither true that one is better than the other nor true that they are of equal value. (Raz, 1986: 322)

To claim that none of the relations "better than" or "equally good" holds between two items is to claim that the items in question are incomparable in respect of their value.<sup>5</sup>

<sup>&</sup>lt;sup>1</sup> See Kuhn (1970).

<sup>&</sup>lt;sup>2</sup> For this distinction, see Lukes (1990), Chang (1997: 1), and Griffin (1997: 39–40), among others. For the view that conceptual incommensurability cannot provide an explanation of moral or political disagreement, see Hurley (1989: 30–54) and Mason (1993: 88–93). In arguing that the lack of common conceptual ground precludes the possibility of substantial disagreement, these authors draw on Davidson's influential critique of the intelligibility of the idea of a conceptual scheme. See Davidson (1986). For an attempt to show that a more modest version of the idea of conceptual incommensurability can be retained as part of an explanation of disagreement (and is even needed to explain deep disagreement), see Richardson (1994: 250–70).

<sup>&</sup>lt;sup>3</sup> Chang (1997) is usually credited with drawing the distinction between these two ideas. Virtually the same distinction is drawn by Stocker (1990: 176–7), though he puts it at a strikingly different use.

<sup>&</sup>lt;sup>4</sup> Compare Stocker (1990: 248ff).

<sup>&</sup>lt;sup>5</sup> It is sometimes claimed that *incommensurability* does not logically entail *incomparability*. For instance, while in his *Morality of Freedom* (1986) Raz takes incommensurability and incomparability to be synonymous, in his *Incommensurability and Agency* (1997) he emphasizes that, strictly speaking, the former does not imply the latter. Raz points out that two paintings whose value is incommensurable can be comparable in a variety of ways: one painting may be bigger than the other, or older etc. (1997: 273, n1). This example suggests that incommensurable items can nevertheless be said to be incomparable given that

Henceforth, I will mainly be concerned with whether it can be established that incomparable items exist. For the sake of simplicity, I will use the terms "incommensurability" and "incomparability" interchangeably.

Before going further, let me set aside other uses of the term "incommensurability". What is sometimes meant by saying that two values are incommensurable is that loss in one value cannot be compensated by any gain in the other value. For instance, one would not give up freedom of movement for any amount of chocolate. But when we say that freedom of movement is incommensurable with the pleasure of eating chocolate, what we actually mean is that is the former is immeasurably more valuable than the latter. Strictly speaking, this is not incomparability, but an extreme form of comparability.<sup>6</sup>

Consider, next, the case of "constitutive incommensurability". Raz (1986: 345–53) argues that one's capacity for entertaining a relationship of friendship depends, *as a conceptual matter*, on judging that friendship is incomparable with money. But this is not incomparability in the sense that concerns us here, because the case that Raz attempts to make (i.e. that friendship and money are incomparable), is not one in which the items involved *cannot* be compared, but rather one in which they *should not* be compared.<sup>7</sup>

#### 4.2 What items can be incommensurable?

What items can be incommensurable? A brief survey of the literature reveals that different philosophers make incommensurability bear upon different items: options or alternatives facing moral agents (e.g. courses of actions or comprehensive goals), as well as values, reasons, moral norms and virtues.<sup>8</sup> Given the diversity of items that are said to be incommensurable, the question that arises is whether these approaches are largely intertranslatable. Offering a comprehensive answer to this question is beyond the scope of this dissertation. However, for the purpose of establishing a working definition of value incommensurability, the following remarks constitute a fair treatment of this issue.

comparisons are at most possible along some dimensions that are ultimately irrelevant for judging which item is more valuable. But this need not be the case. As Raz's analysis of mixed-value goods shows, incommensurable items can sometimes be comparable in several significant respects (see Raz (1991)). For instance, two novels can be compared according to several criteria (originality, insight, linguistic skill etc.), but we might be unable to say which of these criteria should be assigned more weight in case they yield different rankings. Saying that the novels in question are incommensurable comes down to saying that we are unable to form an *overall comparison*.

<sup>&</sup>lt;sup>6</sup> See Griffin (1977: 44–7; 1986: 85–9; 1997: 35–6).

<sup>&</sup>lt;sup>7</sup> For criticizing Raz's view according to which cases of constitutive incommensurability present us with cases of genuine incommensurability, see Reagan (1989) and Chang (2001).

<sup>&</sup>lt;sup>8</sup> Most authors favor an account of incommensurability in terms of values (see Berlin (1959), Nagel (1979), Guttenplan (1979–80), Wiggins (1980), Williams (1981), Raz (1985–86), Seung and Bonevac (1992), Kekes (1993), and Richardson (1994)) or reasons (see Nagel (1979), Raz (1991; 1997), Dancy (1993), and Schaber (2004)). Other authors talk of incommensurability in terms of moral requirements (Sinnott-Armstrong (1988)), morally relevant properties (Shafer-Landau (1994), Gowans (1994)), moral claims (Brink (1996)), moral norms (Griffin (1997)), or virtues (Zangwill (1999)). For a more simplified view, see Schauer (1998: 1215), who defines incommensurability in terms of "values, reasons, options or norms". Schauer uses this phrase to signal his "agnosticism [...] with respect to the contested question of what it is that is commensurable".

(a) We should notice, first, that our interest does not typically lie in comparing in values *per se*, but in comparing options that realize values.<sup>9</sup> Given two values, say liberty and equality, what does the claim that they are incomparable amount to? Insofar as what is meant by it is that we cannot compare liberty and equality in themselves, this claim turns out to be the denial of a hardly meaningful assertion. Raz draws attention to the fact that:

It makes sense to talk of the relative weight of options one can at least in principle choose between. It makes no sense to talk of choosing between perfect liberty and absolute equality. As long as one is a person one has some liberty. Nor is it clear what could be meant by a 'situation of total inequality' [...]. (Raz, 1986: 322)

The upshot of these considerations is that one can meaningfully compare only options that realize to various degrees the value of liberty and the value of equality.

Furthermore, the claim that liberty and equality are incomparable turns out to be plainly false insofar as what is meant by it is that instances of these values can never be compared.<sup>10</sup> At least in some cases, a way of arbitrating between the conflicting demands of liberty and equality is readily available. If an immense sacrifice in liberty were required to bring about a minor increase in equality, it is uncontroversial that liberty would prevail over equality. Much the same holds for other values. To give but one example, although it is often said that knowledge and aesthetic experience are incommensurable, it is obvious that we can compare different instances of these values as more or less valuable. Admittedly, knowledge of the theory of relativity is more valuable than the aesthetic experience gained by watching a mediocre adaptation after one of Molière's plays, whereas knowledge of a minor detail of the British legislation regulating the gaming industry is less valuable than the aesthetic experience gained by regularly attending the Royal Shakespeare Company's performances.

Having discarded the two aforementioned interpretations, consider a third one: what we mean by claiming that liberty and equality are incomparable is that it is neither the case that any liberty-enhancing option is better than any equality-enhancing one, nor that any equality-enhancing option is better than any liberty-enhancing one, and at the same time, it is not the case that any two such options are of equal value. Yet, the claim that two values are incommensurable amounts to rather little if it is only meant to convey the idea that neither of them always takes precedence over the other. Since there are theories that endorse a lexical ordering between values, this is not a trivial point.<sup>11</sup> Nevertheless, the claim of incommensurability is meant to go beyond it. While granting that two given values are not lexically ordered, one can still insist that in any given context one of the relations "better than" or "equally good" will necessarily hold between options realizing the values in question.<sup>12</sup> It is precisely this further claim that advocates of incommensurability want to deny.

<sup>&</sup>lt;sup>9</sup> A sample of authors who explicitly claim that comparisons should be understood as holding between concrete options that instantiate values, and not between abstract values, includes Williams (1981: 77), Raz (1986: 322–3), Chang (1997: 1), Lukes (1997: 184–5), and Broome (1999: 146–7). <sup>10</sup> See, for instance, Williams (1981: 77).

<sup>&</sup>lt;sup>11</sup> See, for instance, Rawls (1971).

<sup>&</sup>lt;sup>12</sup> Think of Griffin's discussion of incommensurability in the prudential sphere. Although he believes that there are irreducibly many prudential values, such as enjoyment, understanding, accomplishment, deep

Where does this leave us? The following definition stands out: two values A and B are commensurable if and only if, for any pair of options a and b that instantiate A and B respectively, it is either the case that a is better than b, or that b is better than a, or else it is the case that a and b are equally good. Conversely, two values A and B are incommensurable if and only if there is (at least) one pair of options a and b such that it is neither the case that a is better than b, nor that b is better than a, and at the same time, it not the case that a and b are equally good.<sup>13</sup>

Adopting the above-mentioned definition places us in a position to reject from the outset an argument often used against incommensurability. According to it, in order to reject the claim that two values are incommensurable, it is sufficient to pinpoint some cases in which comparing instances of these values raises no difficulty. The proponents of this argument believe that incommensurabilists have trouble accommodating such cases because they assume that what incommensurabilists must actually hold is that instances of incommensurable values can never be compared. However, this is a misconstrual of the incommensurabilists' position. Along the same lines, John Broome argues that:

Undoubtedly, *some* amount of pain-avoidance can be compared with *some* amount of accomplishment; indeed there may be no two values such that some amount of one cannot be compared with some amount of the other. [...] But, if pain and accomplishment are to be thoroughly commensurable, *any* amount of pain-avoidance must be comparable with *any* amount of accomplishment, so that all options that realize pain-avoidance or accomplishment to any degree can be ordered by their goodness. (Broome, 1999: 146, emphases in original)

In other words, commensurabilism is rendered false by the existence of a single case in which comparison is beyond our reach, whereas the existence of cases in which comparison is unproblematic cannot render incommensurabilism false.

(b) Although the view that incommensurability should be understood as a failure of comparability among concrete options has gained currency in the recent literature, there are incommensurabilists who still favor the view that incommensurability should be understood in terms of abstract values. Consider, for instance, the following proposal advanced by David Wiggins:

[L]et us rule instead that incommensurables are the relatively unspecific, rather general, potentially conflicting or rival objects of concern that historically determinate options import into the consideration of a given situation. [...] And let us say that A is incommensurable with B [...] if there is no *general* way in which

personal relations and so on, and that these values are not lexically ordered, Griffin doubts that there are any incommensurabilities in the prudential domain. See Griffin (1986; 1991).

<sup>&</sup>lt;sup>13</sup> Alternatively, we can say that two values A and B are incommensurable if and only if there is (at least) one pair of options a and b such that it is neither the case that a is better than b, nor that a is worse than b, and at the same time, it not the case that a and b are equally good. However, nothing important hangs on whether we use one formulation or the other since a's being worse than b implies that b is better than a. Henceforth, I will use these two formulations interchangeably.

A and B trade off in the whole range of situations and choice and comparison in which they figure. (Wiggins, 1997: 58–59, emphasis in original)

In what follows, I will briefly discuss my reasons for adopting the first of the two abovementioned views.

Going back to a point made earlier, reducing the question of whether values are incommensurable to the question of whether we can get lexical or trade-off relations between them renders incommensurability a philosophically less interesting claim than it really is. If, however, the absence of such relations is taken to entail a (partial) failure of comparability among concrete options that instantiate values, this leads to a different problem, namely that incommensurability is tangled up with other difficult issues. As an illustration of this point, consider another passage from Wiggins:

Incommensurability [...] reflects the separateness and mutual irreducibility of the standing concerns that make up our orientation toward the distinct values and commitments (and whatever else) that impinge on us in different sorts of situation. It reflects the fact that these concerns are not variations on a common theme. According to our new account, incommensurability can indeed [...] have *application* to specific historically determined choices or options [...] But it does so only derivatively from the incommensurability (in the new sense) of the more general or persisting concerns to which we have standing, generally unranked attachments. (Wiggins, 1997: 59, emphasis in original)

There are two distinct ideas in this passage. The first has to do with the fact that the values which constitute the objects of our standing concerns are irreducibly plural. The second has to do with the fact that practical choices involving plural values are often underdetermined. Wiggins proposes to use the term "incommensurability" in connection with the first idea, and only derivatively in connection with the second one, since the fact that our practical choices are often underdetermined is explained by the fact that the values which make rival claims upon us are grounded in a plurality of mutually irreducible concerns. While Wiggins might be right in claiming that any sensible account of an alleged failure of comparability among concrete options must rely on pluralism, I believe that we should keep these two ideas separate. It is important to stress that there are authors who argue that the connection between pluralism and incommensurability is not as straightforward as Wiggins seems to imply. On the one hand, there are those who claim that mere pluralism does not entail incommensurability.<sup>14</sup> On the other hand, there are those who contend that value monism can allow for incommensurability.<sup>15</sup> In order to avoid settling such substantive matters at the level of definition, we should reserve the term "incommensurability" for the failure of comparability among concrete options.

<sup>&</sup>lt;sup>14</sup> See, for instance, Griffin (1986: 89–92; 1997: 36).

<sup>&</sup>lt;sup>15</sup> See Chang (1997: 16–7). She argues that it would be a mistake to think that instances of values that can ultimately be reduced to one supervalue cannot be incomparable. Chang points out that sophisticated versions of value monism, which acknowledge qualitative differences among value bearers, can give rise to incommensurabilities. Moreover, she draws attention to the fact that more of one value is not necessarily better with respect to that value. What follows is that different quantities of a single value may be incomparable. The upshot of these considerations is that the issue of monism and pluralism cuts across the issue of commensurability and incommensurability.

However, one can object to this line of thought by pointing out that there are other grounds for claiming that incommensurability should primarily be understood in terms of abstract values. Recall that those who hold that incommensurability should be understood in terms of concrete options stress that comparative judgments cannot *meaningfully* hold among values themselves. On this view, values cannot be ranked "because they come in various amounts" (Broome, 1999: 146). When asked to decide on the relative standing of values themselves, the argument goes, we have trouble understanding what it is exactly that we are required to compare.

By contrast, Wiggins entertains no doubts about the meaningfulness of comparative judgments among values themselves. According to him, the fact that our practical choices are often underdetermined has to do with the fact that the values that make rival claims upon us are grounded in a plurality of concerns to which we have "generally unranked attachments" (Wiggins, 1997: 59). If so, then the question whether values themselves can be ranked is prior to the question whether options realizing them can be ranked. While offering a full-fledged account of Wiggins's view on incommensurability is beyond the scope of this work, a brief point of clarification may nevertheless be useful at this point. Wiggins stresses that the phrase "generally unranked attachments" is intended "to leave room for a distinctive moral emphasis among a rational agent's attachments" (1997: 266, n7, emphasis in original). Yet, even if rational agents arrive at certain emphases in their practical outlooks, this does not mean that the agent's relative ranking of various objects of concern will deliver a general account of how different objects of concern trade off for him. Moreover, Wiggins maintains that, where A and B are incommensurable values or objects of concern, there is no correct account to be had of how A and B trade off against one another. The suggestion here is that the potentially conflicting objects of concern that we constantly seek to promote and accommodate are in a way equally fundamental. Insofar as this picture is coherent, the objection might go, it speaks in favor of an account of incommensurability in terms of abstract values.

A possible way to downplay the importance of this objection is to stress that one can hold that there is a plurality of values that are in a way equally fundamental without contradicting the view that comparative judgments involve options rather than values. Such an approach would proceed by recording the fact that individuals typically display certain concerns. For instance, it can be noticed that all humans show a concern for self-preservation, knowledge, deep personal relations and so on. The next step would consist in identifying these basic concerns and the values corresponding to them.<sup>16</sup> A further step would involve determining whether these values are in a way equally fundamental. One way to do this is by stressing that a good life or an ideal life would have all these ingredients. Inasmuch as a life is poor in personal relations, it is not a good life. This holds regardless of whether other values are realized. Likewise, inasmuch as a life is poor in knowledge, it is not a good life. No comparison with other values is needed in order to establish that knowledge is a fundamental value. Hence, comparative judgments among abstract values are not strictly required in order to establish that there is an irreducible plurality of values that are in way equally fundamental. <sup>17</sup>

<sup>&</sup>lt;sup>16</sup> For such an approach, see Finnis (1980: ch. 4).

<sup>&</sup>lt;sup>17</sup> This argument draws on Stocker's account of pluralism. According to Stocker (1990: 170–1), *lack* is one of the most important marks of plurality. While granting that Stocker is right in claiming that a life poor in knowledge or in sensual pleasure "does not merely not have that good, it lacks it" (1990: 170), and that

However, these remarks leave us well short of holding that values themselves cannot be the objects of comparison. Now, let me turn to an argument put forward by Mozaffar Qizilbash against Broome's view that all comparative judgments involve options that realize certain amounts of value.<sup>18</sup> Qizilbash (2000) considers the following example. Suppose that you have to choose between writing a paper and going punting on the river, and that, in making this choice, you consider accomplishment and enjoyment. Broome would have us believe that in this case comparison is between a certain amount of accomplishment and a certain amount of enjoyment. However, you do not know how good the paper would turn out to be, and thus, how much of an accomplishment it would constitute. Furthermore, though you think punting is fun, you do not know how much you would enjoy it. Qizilbash argues that:

It is possible, in this case, for accomplishment to outvalue enjoyment, without its being clear how much of these values are being realized. [You] might decide to go for writing the paper, simply because [you] think, or decide that, accomplishment is more valuable than enjoyment. So while comparisons of quantities of value can help explain how our rankings of values – in the specific forms they are realized – turn out, they are not necessary for making comparisons of prudential values. Furthermore, judgments about the specific forms in which values manifest themselves can depend on rankings of values in the abstract. (2000: 229)

Some will find this argument unconvincing. For even if the amounts of value that you can realize in choosing one way or the other are subject to uncertainty, you can and do expect that they will fall within a certain range. You might not know quite how much you would enjoy punting, but you can expect that punting is more or less as enjoyable as rowing or other water sports you have practiced. Furthermore, you might not know how good the paper would turn out to be, but based on your experience you can expect that several hours of writing will result in a quite dignified piece of work. Your choice might simply reflect a judgment to the effect that the highest amount of enjoyment that you can hope to realize by going punting is worse than the lowest amount of accomplishment that you expect to realize by writing the paper. Hence, even if the amounts of value that you

only pluralism can make sense of lacks, one can insist that lacks cannot be used to establish that there is a plurality of values that are *equally fundamental*. Reference to the ideal life does not help bolster this argument, for holding that an ideal life is a life in which *various* values are realized is consistent with holding that placing a certain emphasis on *some* of these values is a necessary part of leading an ideal life. Our being able to judge, for instance, that a life rich in knowledge and poor in pleasure is better (i.e. closer to the ideal life) than a life rich in pleasure and poor in knowledge is evidence for this point. Therefore, considerations about the ideal life show that comparative judgments among abstract values surface at a different level. However, it is not at all clear that, when deciding which particular mix of values makes for the ideal life, values in themselves are the relevant objects of comparison. To concede that an ideal life is a life in which different values are realized to different degrees means to concede that comparing instances of values is what we are mainly interested in.

<sup>&</sup>lt;sup>18</sup> Note that Broome's argument to the effect that we cannot meaningfully compare values themselves differs from Raz's argument. While Broome claims that comparative judgments necessarily involve *amounts* of value, Raz's point is rather that, in making such judgments, we cannot abstract from the *specific forms* in which values are realized. (This explains why Raz (1986: 322) doubts that we can meaningfully compare liberty and equality in themselves, but allows that comparisons among liberty and equality make sense if they are understood as "comparisons of types of options", i.e. as comparisons between types of freedom and types of equality.)

can realize are subject to uncertainty, it might still be the case that choice about what option to go for comes down to whether a certain amount of one value is better than a certain amount of another value.

But consider a different example. Suppose you are working for an international company which offers you the opportunity to immediately relocate to a foreign country. You know very little about that country, but this is what you find most challenging about the whole enterprise. You reckon that living abroad is an experience that will definitely contribute to your self-development. However, you realize that, in moving abroad, you must forgo the enjoyments of a life lived in a friendly, familiar environment. Suppose you decide to accept the offer made by your company. This might simply be because you reckon that self-development is more valuable than enjoyment. Note that, in this example, there is little reason to think that choice about what option to go for boils down to a comparison among amounts of value. This reading is ruled out by the fact that – as opposed to the case discussed by Qizilbash - this is a case where uncertainty relates not only to the amounts of value that are being realized, but also to the specific forms in which they are realized. More concretely, when you decide to move abroad, you do not know in what way this experience will contribute to your self-development. Would it help you improve your linguistic skills? Would it increase your sense of toleration by providing you with a first-hand experience of a different cultural environment? Would the job requirements be more demanding? This experience can contribute to your selfdevelopment in *several* ways, some of which you are not even aware of at the moment of making a decision. It might be objected that self-development is, in fact, a placeholder for several values. But this makes it even harder for the proponent of Broome's view to claim that, in this example, comparison is among amounts of value. This is because, although you are confident that *overall* this experience will positively contribute to your selfdevelopment, you are unaware of the precise respect(s) in which you will actually evolve. Therefore, your decision to move abroad does not rely on a comparative judgment among a certain amount of self-development and a certain amount of enjoyment, but rather on a comparative judgment about the relative worth of these values.

The upshot of these considerations is that the claim that comparative judgments cannot meaningfully hold among values themselves is implausibly strong. Now, let me conclude with a few brief observations.

First, even if comparability can hold among values themselves, most comparisons among values should be understood as having a *presumptive* character. That is, a judgment about the relative worth of two values should not be taken to entail that *any* amount of the preferred value outranks *any* amount of the other value, or that there is a statable ratio of substitution between them. (This is, in fact, the position advocated by Wiggins). Examples of the sort discussed above show that we can rank self-development or accomplishment above enjoyment, and that in making such judgments we do not rely on other judgments involving amounts of value. However, holding that accomplishment is more valuable than enjoyment does not commit us to the claim that any instance of accomplishment outweighs any instance of enjoyment. Going back to Qizilbash's example, suppose you had a reason to believe that your paper would turn out to be mediocre (so that it would barely constitute an accomplishment), whereas going punting would be the greatest occasion of enjoyment that you had in years. If so, choosing to go punting would be rational even if you believed that accomplishment is more valuable than enjoyment. In fact, the reason why some people find it so obvious that even the slightest amount of some value outranks even the greatest amount of some other value is that they often take second to be a disvalue. Think of a person who believed that any amount of knowledge is worth more than any amount of sensual pleasure. Undoubtedly, such a belief cannot be dissociated from a practical outlook that confers no value at all to sensual pleasure.

Second, the claim that values as such can be the objects of comparison does not vindicate the view that incommensurability should be understood in terms of values, rather than in terms of concrete options. According to Wiggins, the failure of comparability among concrete options can be traced back to there being no lexical or trade-off relations among values. This view is not rid of difficulties. For instance, Michael Stocker (1990: 200-7) argues that, even though there are not many sound and useful comparisons among values, comparing instances of values is commonplace and often enough unproblematic. Although he does go to some length in defending the second claim, he takes it to convey a fact that is hardly worth denying. This calls for an argument that would clearly establish that there are cases in which practical reason fails to adjudicate between options that instantiate distinct values.<sup>19</sup> Moreover, Stocker stresses that our being able to make sound and useful comparisons among instances of values does not rely on our being able to make sound and useful comparisons among values. In fact, once we acknowledge that there is no general way in which different values trade off against one another, it seems difficult to resist the conclusion that comparative judgments among value instances do not depend on comparative judgments among values in the abstract. If we want to preserve the intuition that whether some objects of concern can be deemed to be more important than others has a bearing on whether options confronting us can be incomparable (an intuition that many incommensurabilists share with Wiggins), we need a more detailed account of how exactly the first issue bears on the second one.<sup>20</sup> In short, my point is that taking Stocker's objection seriously requires an account of incommensurability in terms of options.<sup>21</sup>

(c) There is, however, a further reason for focusing on comparing concrete options, rather than abstract values. As Broome rightly points out:

To ask whether values are commensurable is implicitly to make a presumption about the way individual values combine together to determine the overall goodness of an option. It is to presume that the goodness of an option is determined by the various values it realizes, acting independently of each other. [...] But actually the value of an option may be determined in a complex fashion by the interaction of values with each other and with other features of the option which are not themselves values. (Broome, 1999: 147)

 <sup>&</sup>lt;sup>19</sup> The small-improvement argument is such an argument. I will consider this argument in section 4.4.3.
<sup>20</sup> The argument from multiple rankings can be seen as an attempt to substantiate this intuition. For a discussion of this argument, see sections 4.4.1 and 4.4.2.

<sup>&</sup>lt;sup>21</sup> Whether we conceive of incommensurability in terms values or in terms options that realize values has far-reaching implications. For instance, Broome's view that incommensurability should be construed as vagueness relies on the view that it only makes sense to compare options that realize values. (Broome's concentration on the standard configuration rules out those who wish to claim that it is values themselves that are incomparable. For this suggestion, see Griffin (2000: 286).)

These considerations lead to the conclusion that the question of whether options can be ranked is more general and fundamental than the question of whether values are thoroughly commensurable. Given that choice between alternative courses of action is often only indeterminately guided by appeal to values, it is options that constitute the relevant objects of comparison.

On this view, the value of two options is incommensurable if and only if neither of them is better than the other, nor are they equally good. It is important to emphasize that this view does not imply consequentialism (i.e. the view that the consequences of the available options determine how one ought to act) or teleology (i.e. the view that the goodness of the available options determine how one ought to act). Comparisons of options in terms of greatest value need not proceed in terms of a consequentialist norm. Options can be compared with respect to fulfilling an obligation or conforming to a norm of right conduct.<sup>22</sup>

d) Finally, let us note that some authors favor an account of incommensurability in terms of reasons. Roughly put, reasons are considerations that speak in favor or against options. To say that one has a reason to do X entails that X-ing is valuable, or that acting on that particular reason is conducive to the attainment of a certain valuable end. It is important to stress, however, that talk of reasons and talk of values are largely intertranslatable. Thus, if two sets of reasons are incommensurable, then so is the value of options that they are the only reasons for (and vice versa).<sup>23</sup>

#### 4.3 The covering value requirement

Let me turn now to another preliminary point about incommensurability. It is important to bear in mind that any comparison necessarily proceeds in terms of some "covering value". Chang, who has coined the term, stresses that we cannot sensibly say that one thing is simply better than another; one thing can be better than another only in some respect (1997: 5–7). The term "value" is used in this context in a broader sense than the usual, to denote any consideration with respect to which two alternatives can meaningfully be compared. Chang argues that just as no comparison can be understood without reference to a covering value, judgments of incomparability cannot be understood unless such a value is either made explicit, or at least implied.

If failure of comparability is always relative to a covering value, certain judgments of incomparability will be ruled out as mistaken. These are judgments passed about strikingly different items, such as "laptops and clouds" or "cherries and the laws of thermodynamics". We can start by noting that it makes no sense to say that laptops are better than clouds since there seems to be no consideration with respect to which laptops and clouds can meaningfully be compared. If so, then the claim that they are

<sup>&</sup>lt;sup>22</sup> Broome (1999: 154–5) suggests that we need to grant teleology in order to get a clear view of the relevance of incommensurability. For stressing that comparing options in terms of greatest value can

proceed in terms of a deontic norm, see Raz (1997: 273, n4) and Chang (2002*b*: 665–6, n8). <sup>23</sup> See Raz (1991: 83).

incomparable is best understood as pointing to the lack of a covering value in terms of which comparison can proceed. This, however, constitutes a formal failure of comparability. According to Chang, formal and substantive failures of comparability are two distinct phenomena, and only the latter is relevant for the exercise of practical reason (1997: 27–31). To substantiate this point, Chang argues that practical reason never confronts us with comparisons that fail on formal grounds. Indeed, it is hard to think of a choice that turns on whether laptops are better than clouds, or on whether cherries are better than the laws of thermodynamics. Yet, certain choice situations reveal the existence of a covering value even where, at first glance, one can think of none. For instance, although one might surmise that there is no covering value in terms of which oaks and umbrellas can be compared, during a heavy storm one might want to know whether oaks are better than umbrellas with respect to offering protection from thunder. Apart from pointing to an unapparent covering value, these choice situations help us understand why an alleged failure of comparability becomes relevant only when it is relativized to such a covering value.

Moreover, that often two alternatives admit of reverse rankings can be explained in terms of shifting the covering value. Consider again the example of knowledge and aesthetic experience. If comparison is relative to one's general intellectual abilities, then knowledge of some minor details of the legislation regulating the gaming industry is worth less than the aesthetic experience gained by regularly attending outstanding theatre events. If, however, comparison is relative to one's ability to get a job as a financial inspector, then the former is worth more than the latter.

Nonetheless, putting too much stress on the idea that every comparison is relative to a covering value can be misleading insofar as it suggests that we can entirely dispose of the difficulties raised by comparing any two alternatives by identifying an appropriate covering value. It is surely not the case that judgments of incomparability can typically be traced back to the elusiveness of the covering value. Notably, a good deal of the literature on incommensurability focuses on whether incommensurabilities pervade the moral sphere. Those who maintain that moral agents are often faced with incommensurable alternatives claim that, with respect to *moral goodness*, the merits of these alternatives cannot be compared.

Nor is it the case that we can eschew the difficulties raised by comparing certain alternatives by acknowledging that comparison can proceed in terms of *several* covering values. In our daily experience, it is common that choice between two alternatives can be governed by different values, and that one's choice situation will prompt towards one of these covering values. For instance, a diet based on dairy products and a diet based on citrus have different beneficial effects; if one suffers from ulcer, the first is better than the second, whereas if one has high cholesterol, the second is better than the first. However, acknowledging that one alternative is better in some respect, while another alternative is better with respect to promoting justice, while the other alternative can be better with respect to promoting justice, while the point if what we are after is an answer to the question "What ought I morally to do?" In such cases, what we want to know is which alternative is "better overall", or "better, all things considered".<sup>24</sup>

<sup>&</sup>lt;sup>24</sup> Compare Chang (1997: 31–2; 2004).

Chang acknowledges that even where there is a covering value, it is not necessary that there is comparability with respect to it. Still, she grants too much importance to the covering value requirement. Not only does she overestimate the difficulties in finding an appropriate covering value, but at some point, she mistakes one of the central claims advanced by incommensurabilists for the claim that incommensurability stems from the lack of an appropriate covering value. Chang contends that "the requirement that the putative covering value cover the items is [...] what incomparabilists have in mind when they insist that comparison can succeed only if there is some 'common basis' for comparison" (1997: 29). This is, however, doubtful. Some incommensurabilists argue that the absence of a common measure gives rise to incommensurability. But even if the covering value requirement is satisfied, it may still be possible that there is no common basis for comparison. If we are required to rank the *moral* merits of two alternatives, both deontological and consequentialist considerations may be relevant. However, they are irreducible to a common basis.

### 4.4 Arguments for incommensurability

### 4.4.1 The argument from multiple rankings

I turn now to the argument from multiple rankings.<sup>25</sup> The proponents of this argument hold that two given items are incommensurable if there are multiple legitimate rankings of the items in question and none of these rankings is privileged. More concretely, two given items are incommensurable if the following conditions are met:

- (a) the covering value with respect to which the items in question are compared is a function of *several* criteria;
- (b) the relevant criteria yield *conflicting* rankings of the items in question;
- (c) there is no single correct way of weighing the relevant criteria.

It is important to note that there are two distinct versions of this argument. According to the first version, we may be at loss about how to compare two items with respect to a given value even if the constitutive criteria of the value in question can be ranked against each other. This can happen, for instance, if the verdict yielded by the highest-ranked constitutive criterion conflicts with the verdict yielded by a plurality of lower-ranked constitutive criteria. According to the second version of the argument, incomparability with respect to a given value arises because we face a prior puzzle about how to rank the constitutive criteria of the value in question.<sup>26</sup>

<sup>&</sup>lt;sup>25</sup> Chang (1997) identifies in the literature seven types of arguments for incommensurability: arguments from the diversity of values, arguments from bidirectionality, arguments from calculation, arguments from constitution or norms, arguments from the rational irresolvability of conflict, arguments from multiple rankings, and arguments from small improvements. In this chapter, I discuss the last two types of argument. I will briefly deal with the argument from the diversity of values in the last chapter.

<sup>&</sup>lt;sup>26</sup> The argument from multiple rankings is put forward by Raz (1986; 1991), Sinnott-Armstrong (1988: 68– 70), and Shafer-Landau (1995), among others. Raz (1986: 326) claims that "the most important source of incomparability is 'incomplete' definition of the contribution of criteria to a value". However, certain

The following example constitutes a typical illustration of the first version of the argument. In order to judge which of two authors is more philosophically talented, one needs to compare the authors in question according to several criteria, such as originality, insightfulness, and so on. Now, suppose that author A is more original than author B, whereas B is much more insightful, more clear in expression and more technically proficient than A. Arguably, originality plays the most important role in determining whether a person is philosophically talented. Yet, if the verdict yielded by the highest-ranked dimension of assessment conflicts with the verdict yielded by all other dimensions of assessment, there may be no determinate way of combining these different verdicts into an overall comparison.

Raz (1991) offers a subtle account in support of the claim that our weighing of constitutive criteria of values provides only for limited (i.e. incomplete) ranking possibilities. According to him, some values depend for their existence on sustaining social practices. His argument revolves around the special case of mixed-value goods. Mixed-value goods are constituted by standards which determine ideal combinations of component goods. Consider an example often employed by Raz. The merits of a novel depend on a combination of several distinct goods (e.g. originality, insight, linguistic skill, plotting, and so on). Raz claims that understanding how these component goods interact with one another to create a good novel is part of what it means to fully understand what novels are. Inherent standards of mixed-value goods are, in turn, sustained by social practices.<sup>27</sup> This is where incommensurability comes into picture. For standards of excellence admit of both "fixity" and "flexibility", i.e. while we rely on them for the correctness of our judgments of comparative value, such standards do not establish a complete ranking for all the possible combinations among the component goods of a certain mixed-value good. Raz also emphasizes that:

[J]ust as the existence of valuable options depends on social forms so, up a point, their comparative merits depend on social conventions. [...] Social conventions are contingent and finite. They are exhaustible, and are bound to leave plenty of room for incommensurability. (Raz, 1986: 344)

However, one need not endorse Raz's views on the social dependence of values in order to uphold the incompleteness version of the argument from multiple rankings. All that is needed for this argument to go through is to accept that constitutive criteria of values cannot be subject to cardinal rankings, but at most to ordinal ones. For instance, we can judge that, with respect to philosophical talent, originality is more important than technical proficiency or clarity of thought, without knowing precisely how much more important it is. As we have seen, this may give rise to incomparability.

Let me turn now to the second version of the argument from multiple rankings, according to which incomparability with respect to a given value arises because we face a prior puzzle about how to rank the constitutive criteria of the value in question. This

paragraphs from Raz (2003) can be interpreted as supporting the second version of the argument from multiple rankings. Sinnott-Armstrong and Shafer-Landau explicitly endorse both versions of the argument. <sup>27</sup> According to Raz, the fact that standards of excellence are underpinned by social practices does not entail that they lack objectivity. He stresses that the existence of a social practice does not constitute the 'reason' why a particular mix is superior to another: "The standard which is sustained by the practice, and not the practice itself, is the ultimate direct justification for preferring this particular mix." (Raz, 1991: 96)

allows for two different interpretations. According to the first interpretation, our puzzlement has to do with the fact that there are several legitimate ways of ranking the constitutive criteria of a given value, none of which is privileged. According to the second interpretation, incomparability can be traced back to our principled inability to rank the constitutive criteria of a certain value.

The problem with the first interpretation is that it loads the dice in favor of the incommensurabilist.<sup>28</sup> This is most obvious in the case of morally laden concepts. For, while in the case of mixed-value goods like novels it is easier to admit that there are rival accounts of how different dimensions of assessment should be ranked and that none of them is privileged, opponents of incommensurability will most likely argue that, when it comes to morally laden concepts, things are different. More specifically, they will argue that such concepts afford of a uniquely correct interpretation (i.e. a single correct way of ranking the relevant constitutive criteria).<sup>29</sup>

The second interpretation of the argument raises the following problem. One can claim that an argument that appeals to our principled inability to rank the constitutive criteria of a given value is a peculiar argument for incomparability. Consider a view according to which the following claims are true: (a) episodes of pleasure are intrinsically good; (b) the value of a pleasure episode is determined by its intensity and duration; and (c) in what the comparative value of different episodes of pleasure is concerned, neither intensity nor duration is, as a general matter, more important (i.e. whether we assign greater weight to one dimension or the other depends on how greatly it is exemplified). It should be noticed that accepting claim (c) means to concede that comparing the value of different episodes of pleasure we are faced with a prior puzzle about how to rank the criteria themselves. To sum up, this version of the argument from multiple rankings undercuts the link between our abstract rankings of constitutive criteria of values and our rankings of particular options, thereby undermining the possibility of explaining the latter in terms of the former.

One way out of this difficulty is to claim that constitutive criteria of values have equal relevance and to score options according to each of the relevant criteria. This can be done by assigning weights that are represented by real numbers. On this view, the disagreement between incommensurabilists and their opponents comes down to whether there is always a weighted average that uniquely favors one option.<sup>30</sup> However, this model raises more problems than it solves, since the assumptions on which it relies are questionable. As Broome (1992: 175–6) rightly points out, the formula of weighted average assumes that an improvement in one dimension always counts the same in the

<sup>&</sup>lt;sup>28</sup> One might point out, for instance, that this view sits well with Raz's view about the social dependence of values. Also, one might draw attention to the fact that Raz emphasizes that one of the consequences of his view is the "loosening of the rigid divide between matters of knowledge and matters of taste, between the domain of truth and that of preference" (2003: 57).

<sup>&</sup>lt;sup>29</sup> My main interest is in whether moral objectivists can consistently hold that some moral issues are indeterminate. In this chapter, I attempt to show that doubts about whether there really are cases in which the trichotomy of relations "better than", "worse than" and "equally good" fails are unjustified. However, an argument for incommensurability that relies on the claim that there are several correct interpretations of our evaluative concepts would, for my present purposes, assume too much.

<sup>&</sup>lt;sup>30</sup> For the claim that, in some cases, there is more than one acceptable weighted average, see Sinnott-Armstrong (1988: 69–70). For more on the possibility and limitations of weighted averages, see Levi (1986).

overall assessment of an option, whatever the level of the other relevant dimensions. This is, however, a doubtful move which assumes away organic interactions among different dimensions of assessment. A more plausible move would be to acknowledge that the level reached on a certain dimension will influence differently the overall assessment of an option, depending on the levels reached on the other dimensions involved in the assessment. For instance, an improvement in the typing abilities of a good secretary cannot be just as good as the same improvement in the typing abilities of a bad secretary.<sup>31</sup>

The upshot of all these considerations is that the incompleteness version of the argument from multiple rankings constitutes a more cogent interpretation of this argument. As we have seen, this version relies on the plausible assumption that the constitutive criteria of values are subject at most to ordinal rankings, and not to cardinal ones.

## 4.4.2 Multiple rankings and rough equality

Chang (1997: 22–3) claims that the argument from multiple rankings fails to establish that there are items which are incomparable. She considers an example that constitutes a typical illustration of this argument. Suppose we want to compare Eunice and Janice with respect to philosophical talent. Philosophical talent is a complex covering value (i.e. there are multiple criteria according to which philosophical talent is assessed, such as originality, clarity of thought, insightfulness, and so on). According to some of these criteria, Eunice is better than Janice, while according to others, Eunice is worse than Janice. If there is no single correct way of weighing the verdicts yielded by different criteria – the argument goes – then Eunice and Janice are incomparable with respect to philosophical talent.<sup>32</sup> However, Chang contends that the fact that there are several legitimate ways of ranking Eunice and Janice does not entail that they are incomparable. She asks us to consider a third philosopher, Eunice\*, who differs from Eunice by being only slightly more technically proficient and slightly less clear in expression. Chang goes on to stress that:

[C]learly Eunice and Eunice\* are not incomparable with respect to philosophical talent. How could two things so nearly equal in merit be incomparable? Therefore, if Eunice and Eunice\* are not incomparable on the grounds that they can be multiply ranked, then neither are Eunice and Janice on those grounds. (Chang, 1997: 23)

Some might find the claim that Eunice\* and Eunice are comparable with respect to philosophical talent less than obvious. Undoubtedly, there is a sense in which claiming that two items which are very close in value are incomparable is counterintuitive.

<sup>&</sup>lt;sup>31</sup> On this point, see also Chang (2002*a*: 18ff).

<sup>&</sup>lt;sup>32</sup> Chang does not distinguish between the two above-mentioned versions of the argument from multiple rankings. She takes it for granted that what proponents of this argument must actually hold is that incommensurability has to do with there being no uniquely correct way of ranking the constitutive criteria of values.

However, Chang's argument trades on the ambiguity of "comparability". According to incommensurabilists, two items can be said to be comparable if and only if one of them is better than the other or they are equally good. Once we make it clear that this is the meaning of comparability that is relevant for our purposes – the objection might go – Chang's argument loses its strength. To be sure, that one philosopher differs from another by being just a bit more technically proficient and a bit less clear in expression may give us reason to think that they are, so to speak, in the same league. Nonetheless, if there is no single correct way of weighing the conflicting verdicts yielded by different criteria, then none of the relations "better than", "worse than" and "equally good" holds between them.

However, this reply does not entirely dispose of Chang's argument. According to her, the argument from multiple rankings gives us reason to think that there is more to comparability than the traditional relations "better than", "worse than" and "equally good". The idea underlying her argument is that items that are incomparable with respect to a given covering value cannot differ by more or less with respect to the value in question.<sup>33</sup> If so, then the claim that two items are incomparable seems to be incompatible with the claim that they are roughly equal in value.

It is important to emphasize that this point has not gone unnoticed by incommensurabilists. For instance, Raz (1986) urges us not to dismiss the significance of incommensurability by treating it in terms of rough equality. He is certainly aware of the fact that any account of incommensurability that makes room for rough equality runs into serious difficulties:

[R]ough equality presupposes an ability to establish that the gap between the value of the two options is not great. It presupposes, in other words, or seems to, a comparative judgment of the value of two options. But did we not say that incommensurability is defined as the denial that such comparisons are true of the options concerned? If so then by definition incommensurability is incompatible with rough equality. (Raz, 1986: 331)

However, Raz acknowledges that there are cases when we cannot rank the relevant options as better, worse or equally good, and yet we are right to be indifferent between them (such as, for instance, when we have to choose between a cup of coffee and cup of tee). Such cases lend support to view that incommensurability should be understood in terms of rough equality. Raz's solution to this puzzle is to distinguish between two notions of rough equality. The first such notion, which he rejects as inappropriate for his purposes, presupposes that there is a way of measuring the difference between the value of the relevant options and establishing that it is not great. The second notion has to do with the significance of the choice between the relevant options. According to him, it is the second notion that accounts for the fact that, in some cases, we are inclined to think that the difference between two incommensurable options is great, while in other cases, we are inclined to think that this difference is negligible.

<sup>&</sup>lt;sup>33</sup> This claim is explicit in Chang (2002*a*: 141–3).

## 4.4.3 The small-improvement argument

In this section, I will examine what is usually considered in the literature as the strongest argument for incommensurability, namely the small-improvement argument.<sup>34</sup> Simply put, this argument states that, given two items neither of which is better than the other, a small improvement in one of them does not necessarily make the improved item better than the other. If such a small improvement does not tip the balance, then the two original items are incomparable.

Take an example that is already commonplace in the literature. Suppose, with Raz, that one is faced with a choice between a career as a lawyer and a career as a clarinetist. Suppose, further, that after carefully reflecting on all relevant matters (e.g. one's various talents, the chances of success in each career, the way this choice will affect other things one cares about, and so on), one reaches the conclusion that, with respect to goodness of careers, the career in law is neither better nor worse than the career in music. Some might entertain the suggestion that if neither career is better than the other, then they must be equally good. However, if we slightly improve the career in law, for instance, by adding five dollars to the wage, this would not tip the balance. Therefore, the two original careers cannot be equally good, since improving one of two equal options necessarily makes the improved option better than the other.

Here is another well-known example. Walter Sinnott-Armstrong considers the case of a doctor who must decide whether to operate on a patient. On the one hand, since the operation is very likely to save the patient's life, there is a moral requirement to perform it. On the other hand, since even a successful operation would leave the patient in a state of intense pain, there is a moral requirement not to perform it. In some cases, these moral requirements are comparable, for death is clearly worse than a small amount of pain. However, Sinnott-Armstrong argues that there are also cases when these requirements are incomparable, given that for certain amounts of pain, it is neither the case that death is worse than pain, nor that pain is worse than death, and at the same time, it is not the case that they are equal. Some might insist that, if death is neither better nor worse than some amount of pain, then it must be equal to it. Sinnott-Armstrong emphasizes that, if death were equal to some exact amount of pain, say 1000 units, then a small increase in the amount of pain, say 1001 units, would decide the case. Yet, even if the patient would suffer slightly more pain after the operation, say 1001 units, it might still be true that death is neither better nor worse than pain. Given that equality is a transitive relation, death cannot be equal to both 1000 units of pain and 1001 units of pain. Moreover, there seems to be no reason to think that death equals one amount of pain rather than the other. Consequently, death is incomparable with certain amounts of pain.<sup>35</sup>

<sup>&</sup>lt;sup>34</sup> This argument has been put forward by de Sousa (1974), Raz (1986: 324–6), and Walter Sinnott-Armstrong (1988: 66–8).

<sup>&</sup>lt;sup>35</sup> Note that Sinnott-Armstrong favors an account of incommensurability in terms of *moral requirements*. This is because he is mainly interested in showing that incommensurability provides additional support for the possibility of moral dilemmas, which, in turn, are defined in terms of moral requirements. However, talk about moral requirements can easily be translated into talk about reasons. According to Sinnott-Armstrong, moral requirements are, in fact, moral reasons, although not every moral reason qualifies as a moral requirement. He claims that: "A moral reason to adopt an alternative is a moral requirement if and

The small-improvement argument has the following general form:

- (1) A is neither better nor worse than B (with respect to the covering value C).
- (2) A+ is better than A (with respect to the covering value C).
- (3) A + is not better than B (with respect to the covering value C).
- (4) For any x, y and z, if x and y are equally good and z is better than x, then z is better than *v*. [transitivity of equality]
- (5) A and B are not equally good (with respect to the covering value C). [(2), (3), (4)].
- (6) A and B are incomparable (with respect to the covering value C). [(1), (5)]

There are three main objections to the small-improvement argument. First, one can stress that it is not entirely clear whether the examples used by the proponents of the small-improvement argument are meant to involve mere subjective preferences or judgments of objective value. Insofar as the small-improvement argument is defended by appeal to examples that deal with subjective preferences, it falls short of establishing that objective values, i.e. non-preference-based values, are incommensurable.<sup>36</sup>

According to the second objection, the small-improvement argument cannot yield the conclusion that A and B are incommensurable since premise (1) is unwarranted. Consider, for instance, the following passage from Donald Regan:

[W]e cannot use the 'minor improvement' argument, which we used to clarify the concept of incommensurability, to prove that life as clarinetist and as a livestock farmer are genuinely incommensurable. [...] when I said earlier I had not been able to conclude either was better than the other, that was not at all the same as saying that I had been able to conclude neither was better than the other -I may well have been uncertain about everything: whether the life-as-a-clarinetist was better than the original life-as-a-farmer, or whether it was worse, or whether the two were equal. (Regan, 1989: 1061)

In short, Regan argues that our being unable to judge which of two items is better than the other is insufficient evidence against the claim that one of them is better, for truth may outrun our ability to know it.<sup>37</sup>

A third objection has recently been put forth by Chang. According to her, both proponents and opponents of incommensurability mistakenly assume that the conceptual space of comparability between two items is exhausted by the trichotomy of relations "better than", "worse than" and "equally good". In a number of influential papers, Chang has argued for existence of a fourth positive value relation, which she calls "parity". If Chang is right in claiming that there is a fourth relation of comparability beyond those of

only if it would be morally wrong not to adopt that alternative if there were no moral justification for not adopting it" (1988: 12).

<sup>&</sup>lt;sup>36</sup> This worry is raised by Regan (1989: 1057) as an objection to Raz's use of an example involving a choice between spending the afternoon walking in the park and reading a book at home. A similar worry is raised by Chang (1997: 260, n42; 2002: 668, n12). Chang draws attention to the fact that de Sousa's version of the small-improvement argument is the weakest because it proceeds by appeal to attitudes of preference and indifference. She points out that the argument is stronger if it is cast in terms of rational judgments. According to her, such a strong version of the argument is put forward by Sinnott-Armstrong (1988). <sup>37</sup> For a similar objection, see Sorensen (1991: 298) and Gowans (1994: 54–5).

the traditional trichotomy, then (6) does not follow from (1) and (5). In other words, the small-improvement argument cannot be used to establish the existence of incommensurable items.

### 4.4.4 The epistemic objection

Epistemicists claim that advocates of the small-improvement argument beg the question at issue. The small-improvement argument proves that, given two items A and B, both A and a slightly improved item A+ are likely candidates for being neither better nor worse than B, and hence for being equal to B. Since B cannot be equal to both A and A+, it seems that the only way out of this difficulty is to accept that A and B are incomparable. Nevertheless, epistemicists argue that our being *ignorant* of the value relations that hold between A and B, on the one hand, and A+ and B, on the other hand, is all that is needed to account for such cases. Therefore, unless it can be established that incommensurability provides a better explanation of such cases, it begs the question to insist that A and B are incomparable.

Moreover, epistemicists emphasize that the soundness of the incommensurabilist's position turns on whether we can tell with *certainty* that neither of the considered item is better than the other. Yet, ruling out the relation "better than" is not rationally justified. According to epistemicists, one of the relations "better than" or "equally good" must obtain between any two items. They argue that from the fact that it is difficult to judge which of two items is better than the other it does not follow that the items in question cannot be compared. When confronted with such hard cases – the argument goes – we can only say that we are *uncertain* as to which value relation holds between the relevant items.

Chang (2002b) has pointed out two reasons for thinking that incommensurabilists take the lead in this debate. First, the epistemic view does not do justice to the phenomenology of hard cases. If we consider the phenomenology of such cases, it turns out that the judgment that neither of two considered items is better than the other cannot always be plausibly reinterpreted as a judgment of *uncertainty*.

Chang considers the example of a selection committee within a philosophy department whose task is to appoint one of two candidates for a vacant chair in the department. Suppose that the committee members agree that predicting future merit is inherently difficult, and therefore, they decide to compare the candidates by attempting to determine which of them displays the greatest philosophical talent in his writings. Suppose, further, that after a thorough examination of the candidates' written work, the committee members reach the conclusion that neither of them is better than the other with respect to philosophical talent. In this example, giving due credit to the *expertise* and *thoroughness* of the committee members, and assessing the available information as the *complete* set of information necessary for deciding the case, it is plausible to suppose that there is nothing further to be known about the case that the committee members do not already know. If so, it seems odd to insist that their judgment can simply be traced back to a lack of evidence that either candidate is better. Hence, Chang concludes that, at a

<sup>&</sup>lt;sup>38</sup> See, for instance, Gowans (1994: 55).

minimum, there is some warrant for the judgment that neither candidate is better than the other. And yet it seems rational to judge that if one of them would be slightly better, this would not decide the case.

Epistemicists can reply that, although it may seem that neither candidate is better than other, in fact it is either the case that of them is better, or that they are equally good. On this line of thought, those who judge that neither of two considered items is better than the other commit some kind of cognitive error. However, Chang rightly points out that the systematic occurrence of such judgments gives us little reason to think that an error theory is correct. She argues that the burden of proof falls on those who insist that such judgments are erroneous:

[T]he epistemicist owes us an explanation of where we go wrong in thinking that neither item is better than the other. Perhaps we make different substantive mistakes in different cases [...]. But how then is the fact that we are systematically in error to be explained? [...] Without an adequate explanation of why we systematically make this putative error, we should perhaps be open to the possibility that we make no error at all. (Chang, 2002*b*: 671)

It is worth stressing that those on each side of this debate take their own position as the default position and claim that the burden of proof falls on their opponents. On the one hand, epistemicists argue that the mere fact that we do not know how two items compare is insufficient evidence against the claim that one of the relations "better than", "worse than" or "equally good" holds between them. On their view, incommensurabilists must provide an *independent* argument to the effect that incomparable items exist. The small-improvement argument, which is supposed to be precisely such an independent argument, does not get off the ground since there is no way to establish in the first place that neither of the considered items is better than the other. On the other hand, incommensurabilists respond that if we cannot know which of two items is better than the other, it is hard to see why it *must* be the case that one of them is better, or that they are equally good.<sup>39</sup> According to them, it is the epistemicist's task to pinpoint the error that we make in judging that neither of two considered items is better than the other. It might be thought that epistemicists have the upper hand since they identify such a systematic error. Allegedly, incommensurabilists wrongly assume verificationism, which is a highly questionable account of truth. Yet, if the phenomenology of hard cases shows that there can be positive evidence for the claim that neither of two considered items is better than the other (as opposed to a lack of evidence for the claim that either is better), this charge is unwarranted.

Second, Chang argues that the abstract version of small-improvement argument is not subject to the epistemic objection. She writes that:

The abstract version of the argument avoids the particular worries because it relies not on particular intuitions about whether a given small improvement makes a difference to how certain items compare but on an abstract intuition about whether certain sorts of small improvements could make such a difference. (Chang, 2002*b*: 671)

<sup>&</sup>lt;sup>39</sup> See, for instance, Sinnott-Armstrong (1988: 68).

According to the "abstract intuition", in the case of evaluatively very different items, a small improvement cannot make the difference between an item's being better or worse than another. Take, for instance, two paintings belonging to different artistic movements. Chang stresses that one need to have no familiarity with art in order to think that a small improvement in one painting cannot make a difference to whether that painting is better or worse than another. Or consider two very different careers. It is reasonable to think that a small improvement – such as increasing the wage by one dollar – cannot make the difference between one career's being better or worse than another.

It might be thought that Chang's appeal to the above-mentioned intuition leaves the epistemic objection untouched. The proponent of this objection claims that, even if it is extremely difficult to compare two given items, it is necessarily the case that one of them is better than the other or that they are equally good. There seems to be no inconsistency between this claim and the claim that a small improvement cannot make the difference between one item's being better or worse than the other. The epistemicist can allow that one item – that is, the very same item – is better both before and after the improvement, although we do not *know* which one it is. Going back to the example in which one is faced with a choice between a career in law and a career in music, suppose that the career in law is slightly improved. If the initially considered career in law was better than the career in music, then, obviously, the improvement cannot reverse this ranking. Yet, even if the career in law was worse than the career in music, it might still be true that the improvement does not make a difference to how the careers compare. Once we assume that comparing the two original careers was extremely difficult, there is no reason to think that a slight improvement in one of them can alleviate these difficulties - hence, our continuing ignorance of the careers' comparative value.

However, this reply is unhelpful. The proponent of the epistemic objection holds that one of the relations "better than", "worse than" or "equally good" necessarily obtains between any two items. His commitment to the view that only a small improvement is required in order to effect a switch from an item's being worse than another to its being better follows from his being a strict trichotomist. As an illustration of this point, consider the following example.

Suppose that we judge that a particular career in law (A) is definitely better than a particular career in music (B). Suppose, further, that B is gradually improved until it is definitely better than A.<sup>40</sup> Imagine now that after each minor improvement career B is compared to career A. Presumably, between the B-items that are definitely worse than A and those that are definitely better than A, there will be an "intermediate zone" containing more than one B-item about which we will be *inclined* to judge that it is neither better nor worse than A. While conceding that such an intermediate zone exists, the epistemicist will insist that, although we cannot arrive at a *confident judgment* about which career is better, for each B-item in this zone it must be true either that A is better than B, or that B is better than A, or that they are equally good. Yet, A can be equally good to a single B-item. Let us call this B-item  $B^*$ . By adding fifty cents to  $B^*$ , we get a B-item that is better than A. Given that the difference between these two items is of one dollar, what follows is that the

<sup>&</sup>lt;sup>40</sup> The model used here closely resembles Broome's "standard configuration". Broome's model is described in section 4.6.

transition from B's being worse than A to its being better is the result of a minor improvement.

By contrast, according to incommensurabilists, between the versions of B that are worse than A and those that are better than A, there is an intermediate zone containing more than one version of B for which it is *false* both that A is better than B and that B is better than A, and at the same time, it is *false* that they are equally good. Once we allow that there is an intermediate zone containing *more than one* version of B which is *incomparable* with A (instead of a *single* version of B which is *equal to* A), there is no reason to suppose that this zone is narrow, especially if A and B are very different things. Hence, the incommensurabilist has no trouble accommodating the intuition that a small improvement cannot effect a switch from an item's being worse than another to its being better.

# 4.5 Parity

In what follows, I will consider Chang's argument to the effect that cases of putative incommensurability can plausibly be reinterpreted as cases of parity. As we have already seen, Chang holds that the small-improvement argument proves that there are cases when none of the trichotomy of relations "better than", "worse than" or "equally good" holds. Yet, she offers an interesting argument – called the chaining argument – to the effect that two items which are not related by any of the traditional value relations may nevertheless be comparable. If both arguments are correct, it follows that there must be a fourth relation of comparability beyond those of the traditional trichotomy. Incommensurabilists assume that if it is neither the case that A and B are equally good, then it follows that A and B are incomparable. However, if there is a fourth positive value relation, then the right thing to say about such cases is not that the items are incomparable, but rather that they are "on a par".

The chaining argument runs as follows. Take two items which are not related by any of the traditional relations "better than", "worse than" and "equally good". Suppose, with Chang, that we want to compare Michelangelo and Mozart with respect to artistic creativity. This is a case to which the small-improvement argument applies: it is neither the case that Michelangelo is more creative than Mozart, nor that Mozart is more creative than Michelangelo, and yet it seems that they are not equally creative, since it is rational to judge that a slight improvement in the artistic achievements of one of them would not decide the case.

Now consider Michelangelo. Suppose we can construct a chain of sculptors starting with Michelangelo and ending with a very bad sculptor called Talentlessi, a chain which has the following property. As we move down the chain, each item is only slightly worse than its predecessor in a single respect of creativity. In other words, between any two adjacent items on the chain there is only a small unidimensional difference.<sup>41</sup> Admittedly,

<sup>&</sup>lt;sup>41</sup> Chang acknowledges that the assumption that there is such a continuum connecting Michelangelo and Talentlessi can be challenged. In general, one can argue that, given two items X and Y which are not related by the traditional trichotomy, the sorts of small changes needed to create a continuum connecting X to an

Chang argues, Talentlessi can be compared with Mozart: a very bad sculptor is decidedly worse with respect to artistic creativity than the great musician.

Recall that, according to Chang, the small-improvement argument relies on the intuition that a small improvement in one of two evaluatively very different items cannot make the difference between one item's being better or worse than the other. Chang stresses that the chaining argument rests on the related intuition that two items which are by hypothesis comparable cannot be rendered incomparable by a small unidimensional improvement in or detraction from one of the items. She writes that:

[F]or items that bear very different respects of the covering consideration, incomparability between them cannot be a matter of some small difference in one of the respects borne such that without this small difference the items would be comparable. A small unidimensional difference just does not seem powerful enough to effect a switch from two such items being comparable to their being incomparable. Call this the "Small Unidimensional Difference Principle". (Chang, 2002b: 674)

According to the Small Unidimensional Difference Principle, if Mozart is comparable with Talentlessi, then it follows that he is also be comparable with Talentlessi+, a sculptor who is only slightly better than Talentlessi. Likewise, if Mozart is comparable with Talentlessi+, then he is also comparable with Talentlessi++, a sculptor who is only slightly better than Talentlessi+. Given that a small unidimensional change cannot transform a sculptor who is comparable with Mozart into one who is not, Mozart is comparable with each member of the chain connecting Talentlessi and Michelangelo. Thus, Mozart is comparable with Michelangelo. We have granted that Mozart is neither more creative, nor less creative than Michelangelo, and at the same time, they are not equally creative. By endorsing both this claim and the claim that they are nonetheless comparable, we arrive at the conclusion that there must be a fourth positive value relation that holds between them.

It is important to note that the chaining argument cannot rely on the controversial assumption that comparability is transitive. Let me explain. Most incommensurabilists acknowledge that mere pluralism does not entail incomparability. More concretely, they admit that comparisons among bearers of different values succeed if they involve an exceptionally fine exemplar of one value and an exceptionally poor exemplar of another value. This claim, conjoined with the claim that comparability is transitive, would undermine the incommensurabilist's position by giving way to complete comparability. For instance, if comparability were transitive, from the fact that Michelangelo, Mozart, Dostoievski, Gaudi and Fellini are all better than Talentlessi with respect to artistic creativity, we could straightforwardly conclude that each of the aforementioned artists is comparable with all the others. Likewise, from the fact that a highly successful career in law and a highly successful career in music are both better than a modest career in law

item Xn that is clearly comparable with Y do not qualify as unidimensional changes. This is because a sufficiently large number of small changes in a single respect can trigger a change in a new respect. However, Chang holds that this objection can be met. She points out that: "[W]e can, in at least some cases, add to the predecessors of the item that triggers a new value an inferior quantity or quality of the new value triggered. In this way, the condition that all items on the continuum bear all and only the same values can be preserved in some further cases and each of the small changes in a single respect will qualify as small unidimensional changes." (2002*b*: 678) For a similar suggestion, see Broome (1999: 148).

with respect to goodness of careers, we could conclude that the former two careers are comparable. The same would hold for highly successful careers in engineering, medicine, philosophy or architecture. Moreover, the same argument can be run for a modest career in music and a modest career in law: since both of them are better than a rather unsuccessful career in law, they are comparable. And again, since an unsuccessful career in music and an unsuccessful career in law are both worse than modest career in law, it follows that they too are comparable. Given that for any two items there is, or there can be, a third item such that each of the original two items is better (or worse) than the third, once we assume that comparability is transitive it is difficult to resist the conclusion that any two items are comparable. However, incommensurabilists explicitly claim that:

[M]ore of one thing may be better than a certain amount of another, even if less of the first thing is incommensurate with that amount of the other. (Raz, 1989: 1221, n145)

In other words, incommensurabilists maintain that two items *can* be incomparable even though each of them is comparable to a third item. Therefore, any argument against the incommensurabilist which relies on the claim that comparability is transitive seems to beg the question at issue.

Now, let me return to the chaining argument. If comparability were transitive, from the fact that Michelangelo is comparable with Talentlessi and the fact that Talentlessi is comparable with Mozart, it would follow that Michelangelo is comparable with Mozart. The chaining argument, however, does *not* move from the claims that A is comparable with B and B is comparable with C to the conclusion that A is comparable with C. Instead, it holds that, since A differs from B only slightly in a single respect, and B is comparable with C, A is comparable with C. Chang (2002b: 675) draws attention to the fact that the chaining argument relies on the *weaker* assumption that, at least in some cases, comparability between items that are evaluatively very different is preserved if we make a small unidimensional improvement in or detraction from one of the items.

In what follows, I will argue that Chang fails to establish conclusively that cases of putative incomparability are, in fact, cases of parity. Chang's suggestion is that anyone who accepts that there are items which are not related by any of the traditional relations "better than", "worse than" and "equally good" is bound to accept that such items are nevertheless comparable, for the small-improvement argument and the chaining argument rest on two *related* intuitions. On the one hand, we have the intuition that a small improvement in or detraction from one of two evaluatively very different items cannot make the difference between that item's being better or worse than the other. On the other hand, we have the intuition that a small improvement in or detraction from one of two evaluatively very different items cannot trigger incomparability where before there was comparability. I believe, however, that while accepting the small-improvement argument, one might reasonably wish to reject the chaining argument.

The first thing to be noticed is that accepting the Small Unidimensional Difference Principle leads to counterintuitive results. Consider a modified version of the example of Michelangelo and Mozart. Chang holds that, if we compare Michelangelo's achievements as a sculptor with Mozart's achievements as a musician, we can reasonably conclude that neither of them is better than the other with respect to artistic creativity. Now, suppose there is a chain of musicians starting with Mozart and ending with a very bad musician called Talentlessi\*, a chain which has the property that each successive item is only slightly worse than its predecessor in a single respect of creativity. Surely, Michelangelo can be compared Talentlessi\*: the great sculptor is decidedly better with respect to artistic creativity than a bad musician. Yet, according to the Small Unidimensional Difference Principle, if Michelangelo and Talentlessi\* are comparable, it follows that Michelangelo and Mozart are also comparable. We know, however, that none of the traditional relations "better than", "worse than" or "equally good" holds between them. Therefore, they must be connected by a forth value relation: Michelangelo and Mozart are "on a par".

Thus far, we have only considered Michelangelo's achievements as a sculptor. Yet, we know that Michelangelo was also a great painter. Now, if we compare Michelangelo's achievements as a painter with Mozart's achievements as a musician, it is plausible to claim that neither of them is better than the other with respect to artistic creativity, nor are they equally good. Admittedly, the painter Michelangelo (hereafter Michelangelo P) can be compared with Talentlessi\*. Thus, according to the Small Unidimensional Difference Principle, it follows that Michelangelo P and Mozart are also comparable. It seems, then, that Michelangelo P and Mozart are "on a par", just as Michelangelo S (Michelangelo the sculptor) and Mozart are "on a par". The same argument can be run for Michelangelo A (Michelangelo the architect). We know that Michelangelo's achievements as an architect were impressive, so it is plausible to claim that comparing his architectural achievements with Mozart's musical achievements leads to the conclusion that neither of them is better than the other with respect to artistic creativity. By applying Chang's chaining argument to this case, we reach the conclusion that Michelangelo A and Mozart are also "on a par".

While it seems plausible to maintain that Mozart is "on a par" with Michelangelo S, Michelangelo P, and Michelangelo A, what is problematic is that the chaining argument would lead to the same result if we were interested in comparing Mozart's achievements in music with Michelangelo's achievements in visual arts, namely that they are "on a par". One can certainly argue that, no matter how impressive his achievements in each of the above-mentioned artistic disciplines, Michelangelo would have been deemed less creative if he hadn't been a practitioner of one of these visual arts. Furthermore, one can point out that there is something counterintuitive about claiming both that Mozart is roughly equal in value with Michelangelo S, Michelangelo P, and Michelangelo A, and that Mozart is roughly equal in value with Michelangelo V (Michelangelo the practitioner of different visual arts). It is important to emphasize that rough equality is not a transitive relation.<sup>42</sup> However, it should also be noted that those who hold that cases of putative incomparability can plausibly be reinterpreted as cases of rough equality or parity stress that breakdowns in the transitivity of this relation are rather uncommon.<sup>43</sup> Admittedly, one can consistently claim that A is roughly equal in value with both B and  $B^*$ , where  $B^*$ is slightly better than B. Yet, if  $B^*$  were significantly better than B, we would reject the claim that A can be roughly equal in value with both B and  $B^*$ . The example considered above falls into the second category. Clearly, one cannot hold that Michelangelo V is only slightly better than Michelangelo S, Michelangelo P, or Michelangelo A with respect to artistic creativity. Consequently, the claim that Mozart is roughly equal in value with Michelangelo S, Michelangelo P, Michelangelo A, and Michelangelo V must be rejected. Note, however, that there is nothing counterintuitive about holding that the

<sup>&</sup>lt;sup>42</sup> See Griffin (1986: 97) and Chang (2002*a*: 145).

<sup>&</sup>lt;sup>43</sup> See Griffin (1986: 97).

achievements of Mozart are incomparable with the achievements of Michelangelo S, Michelangelo P, Michelangelo A, or Michelangelo V.

Now consider a second line of criticism. Several commentators have suggested that the weak link in Chang's argument is the Small Unidimensional Difference Principle and that this becomes obvious once we focus on the question of what makes Mozart better than Talentlessi with respect to artistic creativity. For instance, Timothy Macklem (2001: 12) contends that the answer to this question cannot be that Mozart is creative, while Talentlessi is not, because in this case the claim that Mozart is better than Talentlessi would be vacuously true. Macklem draws attention to the fact that "the value of creativity is constituted by a number of different contributory values, and, as it so happens, is established with reference to different considerations in the case of Mozart and in the case of Michelangelo" (2001: 11). He further emphasizes that Chang's aim is to prove that the creative achievements of Mozart and Michelangelo can be commensurated despite their different contributory values. Yet, Macklem points out that from the fact that Mozart and Michelangelo are both more creative than a noncreative person we cannot infer that they are comparable with respect to creativity despite the different nature of their creative achievements. Furthermore, Macklem argues that when we address the question of what makes Mozart better than Talentlessi with respect to creativity, we must set aside all understandings of creativity that are not based on diverse contributory values. As he puts it, showing that Mozart and Talentlessi are comparable with respect to creativity, where creativity is equated with originality and originality is not seen as the product of valuable components, "is not to establish the possibility of commensurability in the face of diverse contributory values" (2001: 11).

In short, Macklem points out that the assertion that Mozart is better than Talentlessi with respect to creativity is a key move in Chang's argument. However, he contends that Chang can rely on this assertion only if it is assumed that Talentlessi is not totally devoid of creativity, and if the claim that Mozart is better than Talentlessi with respect to creativity means that they can be commensurated in terms of different contributory values. Yet, once we make it clear that this is what is needed for the chaining argument to go through, the following problem arises.

Chang rightly emphasizes that "the mere fact that two items are evaluatively very different does not warrant the conclusion that they are incomparable: a 'nominal' bearer of one set of respects can almost always be compared with a 'notable' bearer of another set" (2002b: 677). However, incommensurabilists can admit that notable A is better than nominal B without accepting that notable A can be compared with notable B. As Martijn Boot (2009) points out, the reason why we accept that notable-nominal comparisons are possible is that in such cases there is a significant difference between the amounts of contributory values involved and this suggests that there is no need to make a trade-off between different contributory values. Yet, even if notable A can be viewed as better than nominal B, if B is gradually improved the need to make trade-offs between heterogeneous values becomes manifest. (Macklem (2001: 14-5) suggests that to accept that Mozart is better than Talentlessi with respect to artistic creativity means to accept that it is possible to commensurate heterogeneous values, while this is precisely what has to be proved. Those unconvinced by this claim should consider Boot's argument to the effect that small unidimensional differences trigger the need of a trade-off between heterogeneous values where such a trade-off could initially be avoided.) Boot considers an example in which one has to compare two careers A and B with respect to goodness of careers. While career A has a good salary (10S) and an average working environment (5E), career B has an average salary (5S) and a more pleasant working environment (10E). If we apply Chang's chaining argument to this case, we reach the conclusion that career A and career B are comparable. Boot points out that if we compare nominal career C (1S, 5E) with notable career B (5S, 10E) it is clear that latter is better than the former, since it is better with respect to both salary and working environment. Yet, a series of small unidimensional differences creates a situation in which a trade-off between heterogeneous values cannot be avoided. If career C is gradually improved by small increases in salary, when the salary is larger than 5S, career C is better than career B with respect to salary, whereas career B is better with respect to working environment. Thus, in order to claim that career B and the improved career C are comparable, one has to assume that trade-offs between salary and the quality of working environment are possible, while this is precisely what has to be proved. The upshot of all these considerations is that Chang is wrong in holding that comparability is preserved through a series of small unidimensional differences. Consequently, the claim that cases of putative incommensurability can be reinterpreted as cases of parity must be rejected.

## 4.6 Incommensurability as vagueness

Most incommensurabilists follow Raz in holding that incommensurability should be construed as a determinate failure of the trichotomy of relations "better than", "worse than" and "equally good". On this view, to hold that A and B are incommensurable (with respect to the covering value C) is to hold that it is *false* both that A is better than B and that A is worse than B, and at the same time, it is *false* that A and B are equally good (with respect to the covering value C). However, Broome has recently put forward an argument to the effect incommensurability should rather be construed as an indeterminate failure of the trichotomy of relations "better than", "worse than" and "equally good". According to Broome, the claim that A and B are incommensurable comes down to the claim that, for each positive value relation, it is *neither true nor false* that it holds between them.

In this section, I will attempt to answer the question whether incommensurability should be construed as a determinate failure of the trichotomy of relations "better than", "worse than" and "equally good" or rather as an indeterminate failure. Let me first briefly consider Broome's argument. Suppose that F is an evaluative predicate. Broome asks us to imagine a "standard configuration" for the comparative "Fer *than*". According to him,

A *standard configuration* for a comparative 'Fer than' consists of a chain of things, fully ordered by their Fness and forming a continuum, and a fixed thing called the *standard* that is not itself in the chain. At the top of the chain are things Fer than the standard, and at the bottom things the standard is Fer than. (Broome, 1997: 69, emphases in original)

The main point of contention between an advocate of Raz's view and an advocate of Broome's view concerns the intermediate area between the top and the bottom of this chain. Say  $A_1 \dots A_n$  are the things the chain in the standard configuration consists of, and B is the standard. According to the advocate of the view that incommensurability should be construed as a determinate failure of the traditional trichotomy of value relations, the continuum in Broome's standard configuration would have to be divided in three areas: the top area (i.e. the area which consists of As for which it is true that A is Fer than B), the central area (i.e. the area which consists of As for which it is false that A is Fer than B, false that B is Fer than A, and false that A and B are equally F), and the bottom area (i.e. the area which consists of As for which it is true that B is Fer than A). On this view, there are sharp boundaries between these three areas. It is worth noting that this picture comes close to the picture advocated by the epistemicist, who holds that one of the relations "better than", "worse than" and "equally good" must obtain between any two items. Both the epistemicist and the advocate of Raz's view would have to accept that there are such sharp boundaries. The difference between their positions is that the epistemicist would claim that the central area consists of only one point, i.e. a unique point which is equally F as the standard.

Yet, the problem is that both the aforementioned views violate a well-entrenched intuition of ours, according to which evaluative predicates cannot have sharp boundaries. Let us consider a standard configuration consisting of a chain of painters fully ordered according to their artistic skill and a different painter (the standard) to whom we compare the painters in the chain. Whereas the painters at the top of the chain are determinately better than the standard, those at the bottom of the chain are determinately not better than the standard. It seems, however, the transition from one area to the other cannot be sharp. This is because it is highly unlikely that a painter that is determinately not better than the standard can be transformed into a painter that is determinately better than the standard by a slight improvement of his artistic skill.

Now, those who claim that incommensurability should be construed as determinate failure of the traditional trichotomy might respond that their view can accommodate the intuition that evaluative predicates do not have sharp boundaries. More concretely, they might claim that the boundaries between the three areas on the continuum from Broome's standard configuration can be vague rather than sharp. In other words, the suggestion is that the central area (i.e. the area of incomparability) is bordered by areas of vagueness. On this view, there are five areas on the continuum: an area in which things in the chain are *F*er than the standard; a borderline area in which it is neither true nor false that things in the chain are *F*er than the standard, false that the standard is *F*er than the things in the chain, and false they are equally *F*; a borderline area in which it is neither true nor false that the standard is *F*er than the things in the chain; and an area in which the standard is *F*er than the things in the chain; and an area in which the standard is *F*er than the things in the chain; and an area in which the standard is *F*er than the things in the chain.

However, Broome (1997: 73–4) argues that, although at first glance there seems to be nothing wrong with combining the view that incommensurability should be construed as a determinate failure of the traditional trichotomy with the view that the boundaries of the area of incommensurability are vague rather than sharp, the resulting picture gives rise to contradictions. Let us concede that the advocate of Raz's view is right to suggest that the standard configuration would contain five areas. Now consider any point in the

borderline area neighboring the top area. Broome points out that it is plainly false that the standard is *F*er than this point, since this is false for all the points in the area of incomparability and above. He further argues that, if there is really an area of vagueness, then for any point in this area it is neither true nor false that it is *F*er than the standard. This is because, if it were false that the point is *F*er than the standard, then the point would lie in the area of incomparability. And if it were true that the point is *F*er than the standard, then the point in the upper borderline area, it is false that the standard is *F*er than it, but not false that it is *F*er than the standard. At this point, Broome asks us to apply the following principle:

The collapsing principle, special version. For any x and y, if it is false that y is Fer than x and not false that x is Fer than y, then it is true that x is Fer than y. (Broome, 1997: 74)

According to the collapsing principle, the point in the upper borderline area is *F*er than *B*. As Broome emphasizes, this implies that the point is not in an area of vagueness after all. Thus, we have to conclude that there cannot be such an area of vagueness.

So it seems that any attempt to reconcile the view that incommensurability should be construed as a determinate failure of the trichotomy of value relations and the view that the boundaries of the area of incommensurability are vague leads to a contradiction. Yet, we have seen that the view according to which there are only three sharply bounded areas on the continuum in a standard configuration is deeply counterintuitive. Therefore, Broome concludes that the suggestion that incommensurability involves a determinate failure of the trichotomy "better than", "worse than" and "equally good" must be rejected. The upshot is that incommensurability must be construed as an indeterminate failure of the trichotomy.

However, as Erik Carlson (2004) rightly points out, Broome's argument is subject to a compelling objection. Carlson draws attention to the fact that the collapsing principle is equivalent to the following principle:

*Vagueness symmetry*: It is neither true nor false that x is Fer than y if and only if it is neither true nor false that y is Fer than x. (Carlson, 2004: 96)

Broome (1997: 96) admits that the collapsing principle implies that vague comparisons must be symmetrical with respect to truth-value. He asks us to suppose that some point x is Fer than the standard y, and that it is neither true nor false that y is Fer than x. Broome argues that since x is Fer than y, the asymmetry of the relation "Fer than" implies that y is not Fer than x. Yet, this would contradict our initial supposition, according to which it is neither true nor false that y is Fer than x. Furthermore, suppose that x is not Fer than y, and that it is neither true nor false that y is Fer than x. Broome argues that it is neither true nor false that y is Fer than x. Broome argues that if we apply the collapsing principle to this supposition, it follows that y is Fer than x. This would again contradict our initial supposition, according to which it is neither true nor false that y is Fer than x. Therefore, the collapsing principle is equivalent to vagueness symmetry.

Carlson (2004: 96–7) emphasizes that there are counterexamples to vagueness symmetry. Suppose that we are interested in whether Alf is a better philosopher than Beth. Suppose, further, that Alf and Beth instantiate to the same degree every property

that indubitably contributes to goodness as a philosopher. Yet, Alf has a greater rhetorical skill than Beth. Does this warrant the conclusion that Alf is a better philosopher than Beth? Carlson contends that here may be no definite answer to this question. This is because it may be indeterminate whether rhetorical skill contributes positively to goodness as a philosopher. If so, then it is neither true nor false that Alf is a better philosopher than Beth. Nonetheless, it is clear that rhetorical skill does not contribute negatively to goodness as a philosopher. Consequently, it is definitely false that Beth is a better philosopher than Alf.

Or suppose, with Carlson, that A and A' are two alarm clocks which are alike in every respect, except that A is waterproof. The question whether A is a better alarm clock than A' may have no definite answer, given that it may be indeterminate whether water resistance counts as a good-making property for clocks that are not very likely to come into contact with water. If so, then it is neither true nor false that A is a better alarm clock than A'. Nonetheless, since it is clear that water resistance cannot detract from A'sgoodness as an alarm clock, it is definitely false that A' is a better alarm clock than A.

In short, Carlson maintains that vagueness symmetry is violated by cases involving indeterminately relevant properties, i.e. "properties for which it is indeterminate whether they are positively relevant for an item's goodness (in a certain respect), but definitely false that they are negatively relevant, or vice versa" (2004: 96). Since Broome's collapsing principle is equivalent to vagueness symmetry, such cases give us good reason to reject it.

It is worth stressing, however, that cases involving indeterminately relevant properties are not the only counterexamples to vagueness asymmetry. Carlson (2004: 97–8) asks us to consider a theory claiming that the value of pleasure episodes is determined by their intensity (I) and duration (D). According to this theory, the following claims about the comparative value of different pleasure episodes will hold:

(A) "*x* is better than *y*" is *true* iff  $Ix>Iy \& Dx\geq Dy$ , or  $Ix\geq Iy \& Dx>Dy$ (B) "*x* is better than *y*" is *neither true nor false* iff Ix>Iy & Dy>Dx(C) "*x* is better than *y*" is *false* iff Iy>Ix, or  $Ix=Iy \& Dy\geq Dx$ 

Now, consider two pleasure episodes such that Ix>Iy and Dy>Dx. From (2) it follows that it is neither true nor false that x is better than y. Moreover, from (3) it follows that it is false that y is better than x. Carlson points out that these two claims contradict the collapsing principle. Note, however, that this is a case in which vagueness asymmetry is not due to indeterminately relevant properties. As Carlson emphasizes:

The only two properties involved, intensity and duration, are both determinately relevant. It is just that intensity has, in a sense, greater weight than duration. An episode y's being more intense than an episode x is enough to make it false that x is better than y, whereas y's being longer than x is compatible with its being neither true nor false that x is better than y. Even if the theory expressed by claims (A) to (C) is implausible, it surely seems consistent. It is hard to believe that it contains some *conceptual* mistake. (Carlson, 2004: 98, emphasis in original)

To make this argument more intuitively appealing, consider a paradigmatic vague term like "bald". Comparative predicates associated with this term (i.e. "balder than",

"less bald" and "equally bald") are vague, given that baldness is not merely a function of the number of hairs on one's scalp, but also a function of the distribution of hairs on one's scalp. Questions concerning the exact number of hairs and the exact distribution of hairs that are required for a man to fall within the extension of "bald" are not settled by our linguistic practices. Thus, when it comes to the application of comparative predicates associated with "bald", it is not clear how the two aforementioned factors interact. Now, suppose that we want to compare Harry and Curly with respect to baldness. Suppose that Harry has 2000 hairs evenly distributed on his scalp, while Curly has 20000 hairs distributed at the base of his scalp, so that most of his scalp is left exposed.<sup>44</sup> In this case, it seems that it is neither true nor false that Harry is less bald than Curly, but definitely false that Curly is less bald than Harry. This may well be because, according to our linguistic practices, the following three claims will hold:

- (1) "x is less bald than y" is true iff  $Sx \ge Sy \& Nx > Ny$ , or  $Sx \ge Sy \& Nx \ge Ny$
- (2) "*x* is less bald than *y*" is *neither true nor false* iff Sx>Sy & Ny>Nx
- (3) "*x* is less bald than *y*" is *false* iff Sy > Sx, or  $Sx = Sy \& Ny \ge Nx$

where "N" stands for the number of hairs on one's scalp and "S" for the surface of one's scalp covered by hair.

In other words, when comparing two people with respect to baldness we do seem to think that the distribution of hairs on one's scalp is more important than the number of hairs. However, we cannot say precisely to what extent the former is more important than the latter.

The upshot of all these considerations is that the collapsing principle must be rejected. Given that Broome's argument in favor of construing incommensurability as an indeterminate failure of the trichotomy "better than", "worse than" and "equally good" stands or falls with this principle, it follows that incommensurability should rather be construed as a determinate failure.

<sup>&</sup>lt;sup>44</sup> Here I draw on an example used by Wasserman (2004: 396).
# Chapter 5 Moral disagreement, value incommensurability and moral dilemmas

## 5.1 Ideal agents and moral indeterminacy

In this chapter I will address one of the most powerful objections to the view according to which the argument from disagreement can be refuted by appealing to the idea of moral indeterminacy. Recall that the idea that there are indeterminate moral issues is supposed to account for the possibility of disagreement among agents who suffer from no cognitive shortcoming. This response to the argument from disagreement runs into the following difficulty. Presumably, agents who suffer from no cognitive shortcoming would have to be aware of all morally relevant facts, including those which can render a situation morally indeterminate. And such awareness of the indeterminate character of a moral situation, the objection goes, would undermine the very possibility of disagreement. In this concluding chapter, I will also attempt to tie up a few loose ends, such as whether the objectivist who deals with the argument from disagreement would be better served by construing incommensurability according to the model proposed by Raz or according to the one proposed by Broome. Another loose end concerns the relationship between value incommensurability and moral dilemmas. In what follows, I will attempt to clarify this relationship.

Let me start by considering the objection mentioned above. According to it, it is doubtful that agents who suffer from no cognitive shortcoming would disagree about how to settle a moral issue instead of just agreeing on its indeterminate character. So it seems that the objectivist must either provide a good reason for thinking that agents who suffer from no cognitive shortcoming do not have access to information of the sort that would undermine the possibility of disagreement, or else he must abandon all hope of rebutting the argument from disagreement by appeal to the idea of moral indeterminacy. As Shafer-Landau (1994: 336–9) rightly emphasizes, an easy way out of this difficulty seems to be available to the moral constructivist. Let me detail.

Shafer-Landau draws attention to the fact that, according to moral constructivism, there are no moral truths prior to or independent of the deliverances of some ideal agent or group of agents. What follows is that, on the aforementioned view, there is no way of telling whether a situation is morally indeterminate independently of the deliverances of such agents. In other words, there is no reason to suppose that ideal agents must be aware of the indeterminate character of any given moral situation. Shafer-Landau points out that:

For the constructivist, indeterminacy arises in one of two ways. On a single-agent theory, indeterminacy arises when the idealized agent remains puzzled or suspends judgment about a situation. On a many-agent account, indeterminacy arises when the contractors fail to come to a consensus about the situation or the status of a rule that putatively governs the case. Thus for the constructivist, whether a given case involves indeterminacies is something that can be fixed only after the responses of the relevant idealized agents are in. Facts about the determinacy status of situations are constructed from the responses of idealized agents, and so do not represent the

sorts of considerations that ideal observers can have access to when responding to the world. (Shafer-Landau, 1994: 337)

In short, moral constructivism seems to be uniquely well placed to accommodate the idea that even ideally placed agents can disagree about how to settle a moral issue instead of just agreeing on its indeterminate character. This is because, on a constructivist view, the determinacy status of a given state of affairs is fixed by reference to the deliverances of such ideally placed agents.

However, some might argue that, whereas constructivism may be better suited than realism to answer the objection that concerns us here, it is not entirely clear whether constructivism can offer a convincing reply to the argument from moral disagreement. As already emphasized in Chapter 2, the objectivist is bound to accept that indeterminacy cannot be a pervasive phenomenon. Therefore, it seems that the best strategy available to the objectivist in responding to the argument from disagreement is to stress that there is also a considerable amount of convergence in moral belief. Yet, one can argue that such convergence can be explained without assuming that there are genuine moral truths. In order to put this worry to rest, the objectivist has to show that convergence in moral belief can suitably be regarded as the result of a reliable epistemic mechanism. The question is whether moral constructivism has the resources to dispel this worry. Surely, defending the claim that moral beliefs are produced by a reliable epistemic mechanism seems easier if a version of cognitivism which delivers a stronger form of objectivity is vindicated. Naturalistically minded philosophers might, for instance, insist that defending the above-mentioned claim requires an argument to the effect that moral beliefs are produced by a *causal* mechanism. And whereas moral realists can hope to accommodate such a picture, moral constructivists can certainly not. This is not to say that moral constructivism does not have the resources to show that our moral beliefs are the result of a reliable belief-producing mechanism. However, the burden of proof lies with the constructivist, i.e. it is the constructivist who has to explain why naturalistic constraints on epistemology should be dropped when it comes to the subject matter of morality. (Perhaps the best strategy available to the constructivist is to emphasize that we are willing to drop such constraints in other areas as well. It is worth stressing that the strategy that I have pursued in Chapter 3 was partly based on similar considerations. Yet, given that the argument developed there was intended as a response to naturalistically minded philosophers, and given that such philosophers are typically unwilling to accept arguments which are built around the claim that there is something peculiar about morality which explains why commonly accepted naturalistic constraints fail to apply to it, I have shifted emphasis from what is peculiar to moral beliefs to what is common to moral and factual beliefs. In particular, I have argued that it would be wrong to think that moral facts are causally inert, and therefore, they cannot figure in a causal explanation of our moral beliefs.)

A detailed discussion of the issue of whether constructivism has the resources to show that moral beliefs are the result of a reliable epistemic mechanism would take me too far afield. Instead of addressing this issue, in what follows I will attempt to prove that the realist can reformulate his answer to the argument from disagreement so as to meet the objection that agents who suffer from no cognitive shortcomings would have to be aware of the facts which can render a situation morally indeterminate. Recall that in order to rebut the argument from disagreement one has to show that there is at least one version of moral objectivism which has the resources to account for both (a) the claim that some moral issues are indeterminate and that our failure to convergence on a moral verdict is sometimes due to such indeterminacy, and (b) the claim that there is also a considerable amount of convergence in moral belief and that this convergence is best explained on objectivist assumptions. My main aim throughout this dissertation has been to show that moral realists can adequately defend both the aforementioned claims. In Chapter 2, I have argued that moral realism has no trouble accommodating the idea of indeterminacy. In Chapter 3, I have attempted to prove that the view according to which moral facts are irrelevant to the explanation of our moral beliefs is mistaken, and that, despite common assumptions to the contrary, advocates of both naturalistic and nonnaturalistic versions of moral realism can consistently endorse the claim that moral facts play causal explanatory roles. In what follows, I will argue that the objection discussed above can be met by the moral realist.

According to Shafer-Landau, the most plausible way to meet this objection would be to insist that the crux of the matter is not whether moral objectivists can account for disagreement among agents who suffer from no cognitive shortcomings, but rather whether they can account for *lack of convergence* on a single moral verdict among such agents. I will come back to the issue of whether this is a legitimate move in dealing with the argument from moral disagreement. Assuming, however, that this is in fact the case, it is obvious that considerations about idealized agents and their supposed awareness of the determinacy status of morally laden situations would cease to present a problem for the moral realist. In contrast to the moral constructivist, the moral realist conceives of moral reasoners as deliberating against the background of already existing moral facts. So the realist cannot hold that even ideal reasoners would fail to be aware of the indeterminate character of certain situations on the grounds that at the moment of deliberation there is no such indeterminacy. However, the realist does not have to adopt this line of response. As Shafer-Landau points out, the realist can accept that "ideal reasoners will attain moral unanimity: for determinate situations, they will all agree on what constitutes a uniquely best evaluation; for indeterminate situations, they will all agree that the situation is indeterminate, and that there is no most appropriate assessment to be given" (1994: 343, emphasis in original). On this view, the reason why even ideal reasons can sometimes fail to converge on a single moral verdict is simply that in some cases there may be no such uniquely correct verdict.

The question is whether shifting the focus from whether there can be *disagreement* among agents who suffer from no cognitive shortcomings to whether such agents can sometimes *fail to converge* on a single moral verdict should be allowed to count as a legitimate move in responding to the argument from disagreement. Shafer-Landau argues that this question should be answered in the affirmative. As he puts it, although realists must ultimately deny the prospect of persisting moral disagreement among agents who suffer from no cognitive shortcomings, "this no longer seems so question-begging once we introduce the possibility of moral indeterminacy" (1994: 343). Shafer-Landau rightly emphasizes that those who adopt this line of response to the argument from disagreement must put forward a convincing argument to the effect that moral realism has no trouble accommodating indeterminacy. In Chapter 2 I have attempted to provide such an argument. I believe, however, that there is a further reason for thinking that the realist can

legitimately invoke considerations of the sort discussed above in responding to the argument from disagreement. Let me explain.

Recall from Chapter 2 that the realist's response to the argument from disagreement is not solely based on the idea that some moral issues are indeterminate. As we have seen, realists typically acknowledge that moral indeterminacy is not a pervasive phenomenon, and therefore, they also appeal to other considerations in replying to the argument from disagreement. They point out, for instance, that some cases of moral disagreement reflect the application of shared moral principles under different circumstances, while others can be reduced to disagreements over nonmoral facts. Furthermore, they hold that some cases of moral disagreement can be attributed to the parties' cognitive shortcomings (although it is important to stress that, if the realist adopts a combined strategy of the sort advocated here, the role he will have to assign to cognitive errors in explaining disagreement will be less significant than it is supposed by the proponents of the argument from disagreement). Prejudice, lack of empathy, as well as the emotions and interests that are often associated with moral controversies also figure in the realist's account of why people tend to go on disagreeing. The point is that realists take seriously the challenge raised by the advocates of the argument from disagreement and offer a thorough and nuanced account of why moral disagreement appears to be so widespread and so persistent. It should be kept in mind that the idea that some moral questions admit of no determinate answer is only one of *several* considerations which, taken together, can provide a plausible response to the argument from moral disagreement. We should be suspicious of a victory of the realist won by suggesting that indeterminacy is a pervasive phenomenon, or by maintaining that to dispel the force of the argument from disagreement it suffices to emphasize that moral reasoners may sometimes fail to converge on a single moral verdict. However, it should be clear by now that a different strategy is available to the realist.

Some might object to this line of thought on the following grounds. In Chapter 2, I have claimed that the idea of moral indeterminacy can assist the realist in answering two different versions of the argument from disagreement. And while the realist's response to the *a posteriori* version of the argument may involve several considerations, only one of which has to do with the idea of moral indeterminacy, this does not seem to be the case with the realist's response to the version of the argument which holds that the mere fact that it is possible for fully competent inquirers to disagree about a moral issue gives us reason to reject realism. Advocates of the latter argument maintain that the realist is faced with the following dilemma. On the one hand, the realist can deny that such disagreement is possible, and insist that moral disagreement must involve a cognitive shortcoming. Yet, given the realist's inability to pinpoint the cognitive defect that would explain certain cases of moral disagreement, this move seems to commit the realist to the existence of a specifically moral cognitive ability. If, on the other hand, the realist admits that there is no lurking cognitive defect by reference to which disagreement can be accounted for, it seems difficult to avoid the conclusion that moral truth is beyond our grasp. In Chapter 2 I have argued that, once the idea of moral indeterminacy is in play, the realist can avoid both the aforementioned outcomes. However, one can point out that, insofar as the realist shifts the focus of the discussion to the absence of convergence on a single moral verdict, his response to this specific version of the argument from the disagreement would in fact amount to a flat-out denial that moral disagreement is possible among agents who suffer from no cognitive shortcoming. The question is whether this is a position that the realist would really want to endorse.

To be sure, the realist who maintains that fully informed and fully rational agents would attain unanimity on the determinacy status of moral laden situations does not have to postulate the existence of a specifically cognitive moral ability. If the reason why such agents sometimes fail to converge on a single moral verdict is simply that in some cases there is no such verdict, then we do not need to assume that the parties are *differentially* competent. Therefore, we do not need to explain the absence of convergence on a single moral verdict by reference to the fact that one party is "morally sighted", while the other is "morally blind". Note also that the realist who adopts the above-mentioned strategy can avoid claiming that moral truths are undetectable. As already pointed out, on the present view, the reason why not even the most competent inquirers can know the right answer to certain moral questions is simply that there is no such answer. While this is undoubtedly an elegant way to handle the difficulty presented by the proponents of the argument from disagreement, some might be left with a nagging feeling that there is something wrong with the realist's response. This is because the present version of the argument is not only based on considerations about how a denial of the possibility of a certain kind of moral disagreement would create serious problems for the realist, but it derives part of its appeal from the intuition that such a disagreement is possible in the first place. Thus, even if the realist can show that the "technical difficulties" raised by denving the possibility of fault-free disagreement can be overcome, as long as we believe that such a disagreement is in fact possible we are quite likely to consider the realist's response inadequate.

Shafer-Landau rightly points out that as long as the debate centers on the possibility of disagreement among fully informed and fully rational moral agents it is difficult to tell who gets the upper hand. He stresses that, since neither party in this debate can come up with a fully developed scenario of how the process of moral reasoning among such agents would look like, "we are usually left with a test of intuitions – the noncognitivist thinking it likely that disagreement [among such ideally placed agents] will persist, the objectivist thinking that indeterminacy will explain it away" (1994: 344). Having said all these, I believe that the best line of defense available to the realist is to draw attention to the fact that the argument from moral disagreement rides on the equivocation between two kinds of situations.

Consider a case in which X is confronted with a choice between two incompatible courses of action A and B. Let us suppose that, after careful deliberation, X decides to do A. Now, suppose also that Y is confronted with a similar choice. Yet, even if these two cases are similar in all morally relevant respects, Y decides to do B. This situation comes pretty close to what the advocates of the argument from disagreement have in mind. The intuition they are trying to exploit is that, while a person faced with a choice between two incompatible courses of action A and B may reasonably come to the conclusion that he morally ought to do A, it is not unreasonable for a different person faced with the same choice to conclude that he morally ought to do B. There is, however, a different situation that deserves our attention. So far we have assumed that X and Y are engaged in what we may refer to as first-person moral deliberation (i.e. they are not asked to pass a judgment on what anyone else would be morally required to do under the circumstances, each of them is simply trying to arrive at a first-order moral judgment about what he ought to do).

Now, let us suppose that X' comes to the conclusion that he morally ought to do A and further adds that it would be unreasonable for anyone faced with a similar choice to conclude that he morally ought to do B. Suppose also that Y' comes to the conclusion that he morally ought to do B and that he further adds that it would be unreasonable for anyone faced with a similar choice to conclude that he morally ought to do A. As already pointed out, the proponents of the argument from disagreement start from idea that a certain kind of moral disagreement (i.e. fault-free disagreement) is possible and maintain that it is an advantage of nonobjectivist metaethical views that they can accommodate this kind of moral disagreement. The question is which of the situations described above one typically has in mind when agreeing to the possibility of fault-free disagreement.

Arguably, those of us who tend to agree that it is possible for agents who suffer from no cognitive shortcomings to disagree about a moral issue actually have in mind the first of the above-mentioned situations. For instance, many of us believe that, while one person confronted with a difficult moral choice may reasonably decide on one course of action, it is not unreasonable for another person confronted with a similar moral choice to decide on another course of action. In other words, we seem to believe that two persons confronted with similar choices may decide on different courses of action without any of them committing a cognitive error. It should be emphasized, however, that as long as we believe that this is indeed the case, we *cannot* also believe that, if one of these persons claimed that it would be unreasonable for the other to choose a different course of action, then this person would commit no cognitive error.

In brief, we have seen that the advocates of the argument from disagreement appeal to the intuition that a certain kind of moral disagreement (i.e. fault-free disagreement) is possible. However, once we distinguish between two kinds of situations, it becomes clear that the intuition that the proponents of the argument from disagreement want to exploit is less widespread than it is commonly supposed. Going back to the situations described above, while we must conceive of X' and Y' as *disagreeing* with one another (since each of them would assert his point of view over and against the other's point of view), we can conceive of X and Y as merely *failing to converge* on a uniquely correct moral verdict. Yet, I have pointed out that, while we tend to think of the latter situation as involving no cognitive failure, things are different when it comes to the former situation. The upshot of all these considerations is that there is no reason to view the moral realist's denial of the possibility of *disagreement* among agents who suffer from no cognitive shortcomings as a questionable move in responding to the argument from disagreement.<sup>1</sup>

<sup>&</sup>lt;sup>1</sup> Let me forestall a possible objection. Some might contend that the line of reasoning developed here is objectionable because it leads to a violation of the principle of universalizability. I want to stress, however, that there are at least two arguments to the effect that a person may justifiably reach a conclusion about what he ought to do under certain circumstances without necessarily making a further judgment about what anyone else ought to do. First, one can raise doubts about the idea that two situations can be exactly the same in all morally relevant respects. This line of argument was famously defended by Winch (1972). For a helpful clarification of Winch's view on the limits of the universalizability principle, see Guttenplan (1979–80). He writes that: "[T]he target of Winch's article is the *usefulness* of universalizability in moral reasoning; he never claims that two cases which are known to be *exactly* alike can be resolved differently. His point is simply that, when all the evidence is considered, few if any cases will be alike and this is because the character of the agent is a significant element in our understanding of the concrete reality of moral cases." (Guttenplan , 1979–80: 74, emphases in original) Second, in order to prove that under similar circumstances two persons may justifiably reach different conclusions about what they ought to do, one can invoke considerations about moral dilemmas. More specifically, one can point out that in the case of

One of the advantages of the strategy outlined here (i.e. insisting that agents who suffer from no cognitive shortcoming would sometimes *fail to convergence* on a single moral verdict rather than *disagree* with one another) is that the realist can avoid having to answer a potentially devastating objection. According to it, a person cannot judge that a proposition p is true without also judging that convergence of fully informed and fully rational persons can be expected to occur on that judgment.<sup>2</sup> In other words, one cannot judge that p and, at the same time, judge that somebody else can, without irrationality, refuse to accept that p.<sup>3</sup> It should be noticed, however, that the account developed here implies nothing of the sort. As already emphasized, on this account, agents who suffer from no cognitive shortcomings attain unanimity on the determinacy status of morally

irresolvable moral dilemmas, there are no moral grounds for favoring one of the two incompatible courses of action available to the agent. Thus, one might view choices made in the context of such dilemmas as choices about what kind of person one wants to be, and not necessarily as moral choices. For this line of argument, see Marcus (1980). She writes that: "A frequently quoted remark of E. M. Foster is 'if I had to choose between betraying my country and betraying my friend, I hope I should have the courage to betray my country." [...] But consider a remark of A. B. Worster, 'if I had to choose between betraying my friend, I hope I should have the courage to betray my country and betraying my friend, I hope I should have the courage to betray my country and betraying my friend, I hope I should have the courage to betray my country and betraying my friend, I hope I should have the courage to betray my friend.' [...] Suppose Forster had said that, morally, Worster's position is as valid as his own. That there was no moral reason for generalizing his own choice to all. That there was disagreement between them not about moral principles but rather about the kind of persons they wished to be and the kind of lives they wished to lead. Forster may not want Worster for a friend; a certain possibility of intimacy may be closed to them which perhaps Forster requires in a friend. Worster may see in Forster a sensibility that he does not admire. But there is no reason to suppose that such appraisals are or must be moral appraisals. Not all questions of value are moral questions [...]." (Marcus, 1980: 135–6)

These are two different ways of articulating the intuition that two persons faced with a similar choice may justifiably reach different conclusions about what they ought to do. The first line of reasoning arrives at the conclusion that no two situations can be exactly the same in all moral relevant respects by extending the sphere of what counts as a morally relevant consideration. The second line of reasoning restricts the sphere of morally relevant considerations (perhaps by suggesting that the scope of morality does not extend beyond the realm of duty), and hence, it arrives at the conclusion that in certain cases morality leaves us free to choose which action to perform. The result is quite similar, i.e. on both the aforementioned views, one can arrive at a judgment about what one ought to do under certain circumstances without being committed to any further judgment about what anyone else ought to do. (It should be emphasized, however, that both the aforementioned views provide an elegant solution to the issue of universalizability. On the former view, the principle of universalizability is of no use because the conditions for its applicability are rarely, if ever, met. On the latter view, if morality leaves an agent free to choose which of two incompatible actions to perform, then anyone faced with a similar situation will be morally free to choose whichever alternative).

For my present purposes, it matters which of the two above-mentioned views one adopts to the extent that the former view implies that there is always a determinate answer to the question of what a particular moral agent ought to do under certain circumstances. (If there is such a determinate answer, then surely observers who make a moral evaluation from a third-person perspective need to converge on a single moral verdict providing that they are fully competent.) However, I see no reason for making such an assumption. The fact that, on this view, some specific features that give meaning to the agent's life may be allowed to count as morally relevant considerations is no guarantee of the fact that there will be a uniquely correct moral verdict. As the literature on moral dilemmas shows, certain choices can be so ghastly that no alternative proves better than the other even after taking into consideration such particularizing features.<sup>2</sup> See, for instance, Skorupski (2000).

<sup>&</sup>lt;sup>3</sup> It is worth emphasizing that the idea that a person cannot judge a proposition to be true without incurring a convergence commitment is an expression of a general concern with the *rationality* of our judgments. So, taking into account the fact that some might have doubts about whether the universalizability principle is a logical principle (as opposed to a substantive moral principle), it can be argued that the objection discussed here is not just a restatement of the objection at *supra* note 1.

laden situations (i.e. when the situation is determinate, they converge on a single moral verdict; when the situation is indeterminate, they agree that there is no uniquely correct moral verdict). What about the suggestion that a person can justifiably reach a decision about what he ought to do under certain circumstances without being committed to any further judgment about what anyone else ought to do? As it will shortly become apparent, as long as we confine the argument to a certain class of first-order moral judgments, we can avoid having to deal with the above-mentioned objection. More concretely, once it is made clear that the argument is intended to apply only to cases in which there is a conflict between two equally stringent moral oughts, this objection can be avoided. (At this point, it is worth recalling that such conflicts are not to be conceived of as conflicts between two all-things-considered moral oughts, but merely as conflicts between two competing non-overridden moral oughts. As I have pointed out in Chapter 2, even though in such cases one has an all-things-considered obligation to adopt either one course of action or the other, the non-overridden moral obligation that one chooses to disregard retains some moral force). An agent who suffers from no cognitive shortcomings must be aware of the fact that he has an all-things-considered obligation to adopt one of the two incompatible courses of action open to him, and that, whichever course of action he chooses, he will offend against a *pro tanto* obligation. Moreover, he must be aware of the fact that, faced with a similar dilemma, another agent can choose differently without being guilty of any cognitive error. To believe otherwise would mean to believe that in the dilemma at issue one has an all-things-considered obligation to adopt a particular course of action, which would amount to a cognitive error.

This brings me to another issue announced at the beginning of this chapter, namely whether the realist who deals with the argument from moral disagreement would be better served by construing value incommensurability according to the model proposed by Raz or according to the one proposed by Broome. Recall that in Chapter 4 I have argued that the former model is preferable to the latter one. Now consider the following objection.

The bulk of this dissertation has been devoted to analyzing the suggestion that the idea of moral indeterminacy can help the realist rebut the argument from disagreement. In Chapter 2 I have distinguished between alethic and comparative indeterminacy and I have argued that neither of them poses a threat to moral realism. In Chapter 4 I have attempted to prove that there are cases when the trichotomy of relations "better than", "worse than" and "equally good" fails. Yet, some might contend that, insofar as my aim was to show that lack of moral consensus can be traced back to the fact that some moral issues are indeterminate, insisting that incommensurability is to be conceived of as a determinate failure of the traditional trichotomy "better than", "worse than" and "equally good" seems problematic. First, a terminological point: some might argue that claiming that in cases of comparative assessment the truth-values of all the relevant statements are determinately false makes the use of the term "indeterminacy" is inadequate. It is important to notice, however, that those who pursue the strategy analyzed here are not oblivious to this point. For instance, Shafer-Landau (1995: 89) maintains that comparative moral indeterminacy arises when the traditional trichotomy of value relations fails determinately. Nonetheless, he stresses that, on his account, a situation counts as determinate if and only if (a) it can be described by a proposition that is either true or false, or (b) it involves a comparison that yields a single best option (1994: 343).

Second, some might insist that the issue is not merely terminological. The worry is that the view according to which the failure of the traditional trichotomy of value relations must be understood as an indeterminate failure may actually be better suited to explain disagreement among agents who suffer from no cognitive shortcomings. Supposedly, if each party singles out one option as the best, and if the trichotomy fails indeterminately (that is to say, if it is neither determinately true nor determinately false that one option is better than the other), then there is a sense in which both parties can be said to be right, whereas if the trichotomy fails determinately (that is to say, if it is determinately false that one option is better than the other), then we should rather conclude that both parties are mistaken.

We are now in a position to see why this line of reasoning is mistaken. Suppose that X is confronted with a choice between two incompatible courses of action A and B, and that, after careful deliberation, X decides to do A. Suppose further that Y is confronted with a similar choice and that Y decides to do B. Now, let us assume that this is a choice involving incommensurable options and that moral reasons on each side are sufficiently strong to yield an obligation. The question is in what sense can X and Y be said to be right? Surely, if X claims that he has an all-things-considered obligation to do A, then he is mistaken. And he is mistaken regardless of whether we interpret incommensurability as a determinate failure of the standard trichotomy of value relations or as an indeterminate one. (X would have an all-things-considered obligation to A if one obligation overrode the other, a condition which does not obtain on any interpretation.) If, on the other hand, Xacknowledges that he has an all-things-considered obligation to do either A or B, and that, while he chooses to do A, somebody else in that situation may reasonably choose to do B, he is certainly right. This is precisely why if we want to hold on to the idea that X and Y suffer from no cognitive shortcomings, we should not think of them as disagreeing with one another, but simply as failing to converge on a single moral verdict.

## 5.2 Value incommensurability and moral dilemmas

In Chapter 2, I have suggested that the realist who wants to establish that comparative indeterminacy poses no threat to moral realism can appeal to considerations about moral dilemmas. More specifically, I have stressed that some might argue that the lack of a right answer to questions of comparative moral assessment is inimical to realism not because it points to an indeterminacy about whether a particular property is instantiated in a given case, but because it threatens to lead to inconsistencies. Yet, the realist can reply that that the appearance of inconsistency arises only on a particular interpretation of the oughts that are in place in cases of comparative indeterminacy. Such oughts yield a contradiction only if we interpret them as all-things-considered moral oughts. However, as it is often emphasized in the literature on moral dilemmas, there is no reason to conceive of such conflicting oughts as all-things-considered moral oughts.

Whereas in Chapter 2 my aim was to prove that comparative indeterminacy poses no serious threat to moral realism, in Chapter 4 I attempted to show that there are cases when the standard trichotomy of relations "better than", "worse than" and "equally good" fails. As I have repeatedly emphasized, we should view metaethical issues as continuous with ethical issues. In particular, I have argued that it is doubtful that the issue of whether the realist can offer an adequate reply to the argument from disagreement by appealing to the idea of moral indeterminacy can be treated independently of other issues (such as the issue of whether there are cases when one is confronted with morally incommensurable alternatives, or the issue of whether a conflict between two non-overridden moral oughts exhibits some inconsistency). Yet, despite suggesting that the realist who responds to the argument from disagreement can rely on some of the arguments advanced by friends of incommensurability and dilemmas, so far I did not clarify the relationship between these two notions. And given that several common assertions about this relationship have been the object of intense criticism, attempting such a clarification would be a good idea.

Before going on, however, let me mention two other reasons for thinking that the debate on moral dilemmas has a bearing on the issue that constitutes the primary focus of this dissertation, namely whether the realist can refute the argument from disagreement. First, as we have already seen, the realist has to prove that there are cases of comparative moral indeterminacy. Yet, some might attempt to resist this conclusion by insisting that, while arguments in favor of incommensurability may establish the existence of marginal, small-scale incommensurabilities, they fail to establish the existence of what Raz calls "significant incommensurabilities". As Raz rightly emphasizes, some might reply to the argument from small improvements that if we make one option not just perceptibly, but significantly better, this will be sufficient to make the improved option better than the other. In other words, some might insist that breakdowns in comparability are "marginal cases, bounded by the test of significance" (Raz, 1986: 327). Moreover, some might raise doubts about whether from the fact that in the case of trivial choices our options might prove to be incommensurable we can infer the existence of moral incommensurables.

However, the literature on moral dilemmas abounds with examples that suggest that such doubts are unwarranted. Consider a well-known example that comes from Bernard Williams (1973*c*: 98–100). While being on a tour through South America, Jim is offered a choice between killing one Indian, thus having another nineteen Indians released, and refraining from killing at the cost of witnessing the murder of twenty men. Now, let us grant that in this example neither choice is better than the other. Let us also modify the example so that one of the options would be significantly better. Suppose, for instance, that by killing one Indian Jim would be able to save the lives of another thirty innocent Indians. Would this be sufficient to make the improved option better than the other? Certainly not. Thus, insofar as we grant that in the above-mentioned example there are reasons to deny that one option is better than the other, the view according to which there are no "significant incommensurabilities" seems unwarranted. To put it more clearly, the suggestion made here is that a careful consideration of some of the most vivid examples in the literature on moral dilemmas would reinforce the idea that it is not the case that our options might prove to be incommensurable only in the case of trivial choices.

At this point, it is worth considering a remark made by Ronald Dworkin. According to him, claims of indeterminacy need more argument than advocates of this view usually provide. Dworkin holds that:

[I]ndeterminacy is a substantive view to be ranked alongside the other substantive views in the neighborhood – that Picasso was greater than Beethoven or *vice versa*, for example. We cannot demand any greater character or level of demonstration for

either of the two latter substantive views than we demand for the former. (Dworkin, 1996a: 5)

Dworkin stresses that those who believe that certain propositions are indeterminate seem to forget that the default view is not *indeterminacy*, but merely *uncertainty*. He points out that we do need positive reasons in order to think that a certain evaluative proposition is indeterminate. Dworkin argues that, for instance, whether there is a right answer to what one ought to do in a moral dilemma will depend "on complex, highly theoretical issues of substantive morality" (ibid.). Admittedly, those of us who are drawn to utilitarianism will believe that there is such a right answer. Thus, in order to conclude that some of the most frequently discussed moral quandaries admit of no right answer, we must have rejected utilitarianism.

I believe, however, that the suggestion according to which it is an apparent neglect of issues of substantive morality that leads some philosophers to conclude that there are cases of comparative moral indeterminacy is misleading. It should be stressed that, while the literature on value incommensurability is replete with over technical arguments, most writings on moral dilemmas display an undeniable concern for substantive moral issues. For instance, Williams's case of Jim and the Indians is intended to expose a major flaw in utilitarianism. The problem is that utilitarianism fails to account for the difficulties that a moral agent would experience in such a case. Williams insists that the resolution of Jim's moral dilemma would require taking into consideration several complex issues, such as "the distinction between my killing someone, and its coming about because of what I do that someone else kills them" (1973c: 117). On his view, the problem with utilitarianism is not only that it suggests that there are *obvious* answers to the question of what one morally ought to do in such cases, but also that "it cuts out a kind of consideration which for some others makes a difference to what they feel about such cases: a consideration involving the idea [...] that each of us is specially responsible for what he does, rather than for what other people do" (1973c: 99, emphasis in original). Therefore, utilitarianism makes the concept of integrity unintelligible. Whether or not one agrees with Williams in that the assuredness with which utilitarianism treats complex moral issues is to be viewed as a flaw in this theory, it is quite clear that holding the view that cases like the one discussed above are genuinely dilemmatic does not have to be explained by a failure to engage in substantive moral argument.

A second reason why the debate on moral dilemmas is relevant for my present topic is that both proponents and opponents of the idea of moral indeterminacy emphasize that the issue of whether two persons can fail to converge on a single moral verdict without any of them committing a cognitive error turns on the issue of whether there are cases when there is no right answer to the question of what one morally ought to do.<sup>4</sup> Dworkin, for instance, who is one of the most fervent critics of the "no right answer thesis", draws attention to the fact that many philosophers erroneously suppose that indeterminacy is the default position in morality and argues that:

There are, in fact, two versions of this default thesis: a first-personal version, which assumes that a judgment of indeterminacy is the right one for someone to reach in the course of his own, personal, first-level moral experience when he is deeply troubled by some moral issue, and a third-personal version, which assumes that a

<sup>&</sup>lt;sup>4</sup> See, for instance, Williams (1981: 72–3).

judgment of indeterminacy is the right one for external observers to reach when they find that other people disagree in the personal, first-level moral judgments they make. I shall discuss mainly the internal version of the thesis, because once we see why the default thesis fails in that version, we see why it falls in the other as well. (Dworkin, 1996*a*: 1)

So, it seems that focusing on moral dilemmas would be a straightforward way of settling the issue of whether agents who suffer from no cognitive shortcoming can sometimes fail to converge on a single moral verdict. (In the previous section I have shown in detail how considerations about moral dilemmas can be brought to bear on the present issue.) Yet, a question that deserves more attention than I am able to devote to it concerns the reasons that have led many philosophers to think that looking at one-person moral conflict affords a better insight into the issue of whether moral conflicts must always be capable of being rationally resolved. For instance, it seems plausible to suppose that, as far as some of the most common examples of "moral dilemmas" go, the question of whether moral conflict admits of rational resolution is not complicated by other issues with which the question of whether interpersonal moral conflicts can often be traced back to the fact that the relevant moral values are interpreted differently by the parties, at least some of the moral dilemmas that have been described in the literature can arise independently of interpretive questions about value).<sup>5</sup>

In brief, dealing with moral dilemmas would provide a response to those that follow Dworkin (1996*a*: 7) in thinking that the view according to which some moral issues are indeterminate is merely the result of transplanting a justification designed for one domain (e.g. a justification designed to convince us that there is some indeterminacy in matters of artistic opinion) to another domain (i.e. the moral domain). Moreover, a careful analysis of one-person moral conflict seems to offer a straightforward way of settling the issue of whether agents who suffer from no cognitive shortcoming can sometimes fail to converge on a single moral verdict.

### 5.3 Defining moral dilemmas

I turn now to the relationship between incommensurability and moral dilemmas. While some authors maintain that incommensurability plays an essential role in explaining the possibility and nature of moral dilemmas<sup>6</sup>, others insist that incommensurability does no explanatory work<sup>7</sup>. The remainder of this chapter is devoted to clarifying the relationship between incommensurability and moral dilemmas.

<sup>&</sup>lt;sup>5</sup> See also Williams (1981: 73), for stressing that "the type of one-person conflict which has in fact been most studied – the so-called conflict of obligation" constitutes "the area of the conflict of values which is most directly linked to reasons for action". Williams also points out that: "Very many of our conflicts [...] including those that have most interested Berlin, are at a level where interpretation in action is less determinate or immediate. Values such as liberty, equality, and expressions of justice other than equality, can certainly conflict as ideals or objectives, though their connection with immediately presented courses of action may often be problematic [...]." (1981: 75)

<sup>&</sup>lt;sup>6</sup> See, for instance, Raz (1986: 357–66) and Richardson (1994: 115–6).

<sup>&</sup>lt;sup>7</sup> See, for instance, Conee (1989), McConnell (1993), and Zangwill (1999: 80).

As already pointed out in Chapter 2, the precise definition of moral dilemmas is a matter of philosophical controversy. In Chapter 2 my aim was not to provide a definition of moral dilemmas, but simply to warn against the mistaken assumption that if there is to be any substantial disagreement between proponents and opponents of moral dilemmas, then dilemmas must be defined as cases in which there is a conflict between all-thingsconsidered moral oughts. I have argued that we do not need to attribute to proponents of dilemmas the view that conflicts between all-things-considered moral oughts are possible in order to conclude that something substantial is at issue in this debate. More concretely, I have argued that, if we look at how different authors deal with the question of whether the resolution of moral conflicts leaves a moral residue, a clear point of contrast emerges between proponents and opponents of dilemmas. Whereas friends of dilemmas typically hold that the resolution of moral conflicts leaves a moral residue, foes of dilemmas reject this view. Friends of dilemmas argue that it would be a mistake to think that overridden obligations are annulled. According to them, such obligations exert some moral force, which gives rise to residual obligations. By contrast, foes of dilemmas typically insist that overridden obligations have no actual weight. Some authors even go as far as to deny that conflicts of obligations are possible. For instance, Alan Donagan maintains that the prima *facie* obligation with the lesser potential weight is not an actual obligation at all, and that "not only do you have no conflict of duties, you do not even have the ghost of a conflict" (1996: 20). When it comes to cases in which there are two non-overridden obligations. foes of dilemmas claim that morality mandates a disjunctive solution, i.e. that all things considered one ought to adopt either one course of action or the other. Yet, they point out that acting on the all-things-considered ought leaves no moral residue, and that the type of practical conflict that arises in such cases should not be confused with a moral conflict. Friends of dilemmas, on the other hand, argue that in such cases there is a genuine moral conflict, and that even though all things considered one ought to adopt either one course of action or the other, the original ought that one chooses to disregard retains some moral force.

At this juncture, a brief point of clarification is in order. By claiming that a major point of contention between friends and foes of dilemmas is whether moral conflicts are soluble without a remainder I do not mean to imply that an adequate definition of moral dilemmas can be provided solely by reference to the idea of moral residue.<sup>8</sup> It should be emphasized that, while maintaining that the resolution of moral conflicts leaves a moral residue, one can also claim that it is the possibility of conflict between non-overridden moral oughts that raises philosophically interesting questions, and therefore, the term "moral dilemma" should be reserved for such cases.<sup>9</sup> Or, one can argue that this term should be used to refer to tragic choices, when there is a great moral residue and when all paths to eliminating this residue are blocked.<sup>10</sup> In what follows, however, I will not enter the debate about how to define moral dilemmas. Undoubtedly, any attempt to clarify the

<sup>&</sup>lt;sup>8</sup> Donagan argues that, in order to transform every case of moral conflict into a moral dilemma, "*no more is needed than to consider prima facie duties as not potential but actual, although capable of being overridden*" (1996: 20, emphasis in original). Yet, as long as we do not define moral dilemmas solely by reference to the idea of moral residue, this is not the case.

<sup>&</sup>lt;sup>9</sup> See, for instance, Sinnott-Armstrong (1988: 19–21, 97–102). See also Nagel (1979: 128–9). Compare Williams (1973*a*), Raz (1986), and Marcus (1996).

<sup>&</sup>lt;sup>10</sup> See, for instance, Raz (1986).

relationship between incommensurability and dilemmas would require at least a working definition of moral dilemmas. Yet, instead of going into the intricacies of the debate over the proper definition of moral dilemmas, I will identify several claims that are endorsed by most proponents of dilemmas, and then I will argue that, regardless of how we define moral dilemmas, incommensurability will play an important role in accounting for them.

Moral dilemmas are sometimes defined as situations in which an agent ought to do *A*, ought to do *B*, can do each, but cannot do both *A* and *B*. Consider a commonplace example. Jean-Paul Sartre (1956) tells of a French student whose elder brother had been killed in the German offensive of 1940. The student reckons that he has compelling moral reasons to join the Free French and fight the forces he regards as evil. However, he realizes that his old mother cannot do without him, nor face the prospect of his potential death. Hence, he also has compelling moral reasons to stay home and take care of his mother. Since he both ought to join the military and ought to attend to his mother's needs, but cannot do both, he seems to be faced with a moral dilemma. Consider another widely discussed example that comes from William Styron (1980). In Styron's novel, Sophie and her two children are on their way to a Nazi concentration camp. A malevolent guard asks Sophie to choose which of her children shall be killed, so that the other shall be allowed to live. In this tragic choice, Sophie has the moral obligation to save each child. Although she can carry out each obligation, she cannot comply with both.

Christopher Gowans (1994: 4-6) has pointed out that defining moral dilemmas as cases in which compelling moral considerations favor each of two incompatible courses of action available to an agent obscures the fact that the debate about moral dilemmas has concerned two distinct issues. The first is an inquiry into the possibility of irresolvable moral conflicts, whereas the second is an inquiry into the possibility of inescapable moral wrongdoing. Much of the debate about moral dilemmas has focused on the question whether there are situations in which there are no moral grounds for choosing between two incompatible courses of action, each of which is supported by some moral reasons. Gowans argues that, once the previous question is answered in the affirmative, a further question arises as to whether agents faced with such situations are condemned to moral failure no matter what they do. He rightly points out that the issue of irresolvable moral conflicts and the issue of inescapable moral wrongdoing are independent of one another. First, the fact that there are cases of moral conflict in which the moral reasons in favor of taking one course of action do not override the moral reasons in favor of taking the other course of action does not entail that moral wrongdoing is unavoidable. Further argument is needed to establish that in such cases, whatever the agent does, he will do something wrong. Second, inescapable moral wrongdoing does not seem to presuppose the existence of irresolvable moral conflicts. Gowans draws attention to the fact that those who claim that there are cases when one cannot escape wrongdoing typically insist that this might be the case even when the moral reasons in favor of taking one course of action override the moral reasons in favor of taking the other available course of action.

Gowans is certainty right to point out that the debate about moral dilemmas has concerned distinct issues, and that some authors use the term "moral dilemma" to refer to irresolvable moral conflicts, while others use it to refer to cases when one cannot escape wrongdoing. He is also right to stress that the possibility of irresolvable moral conflicts is less controversial than the possibility of inescapable moral wrongdoing. Although some of the opponents of moral dilemmas concede that there are certain – rather exceptional –

circumstances in which there is no moral ground for choosing between two incompatible courses of action, each of which is supported by some moral reasons, they categorically oppose the idea of unavoidable wrongdoing. However, if we attribute to proponents of moral dilemmas the view that in some cases one cannot escape wrongdoing, we need to be clear about the sense in which the term "wrongdoing" is used.

Three points are worth emphasizing here. First, as we have already seen, friends of dilemmas share the view that the resolution of moral conflicts leaves a moral residue. It is important to note, however, that there is more than one way to fill out the notion of moral residue.<sup>11</sup> One can argue, for instance, that overridden obligations still exert some moral force, which gives rise to residual obligations. (The same would hold for non-overridden obligations which are not acted upon because of a conflict between two non-overridden obligations). Much of the debate about moral dilemmas has been concerned with whether we owe explanations, apologies, or compensations to someone whose interests have been neglected, and whether we should conceive of such obligations as being generated by the moral force of the overridden obligations rather than by certain fundamental values, such as "minimizing unhappiness and showing respect for persons" (McConnell, 1996: 44). A different issue that has been debated under the heading of "moral residues" is whether there are cases in which an agent would be justified in feeling regret, guilt, or remorse, no matter what he did. While advocates of moral dilemmas agree that overridden obligations can generate residual obligations<sup>12</sup>, there seems to be less agreement among them when it comes to discussing the appropriateness of certain feelings in the aftermath of acting in a moral dilemma. (Some of these divergences might be removed by clarifying the meaning of the terms involved. For instance, if guilt or remorse is supposed to imply a belief that one has acted against an all-things-considered moral ought, then it is doubtful that there are cases in which an agent would be justified in feeling guilty or remorseful no matter what he did. Yet, advocates of dilemmas typically hold that some form of moral distress is appropriate when acting in a moral dilemma. And it seems plausible to assume that the degree of moral distress varies according to circumstances.<sup>13</sup>)

Second, the term "wrong" is used somewhat differently by different authors. For instance, Walter Sinnott-Armstrong writes that "it is not morally wrong to violate either moral requirement in a moral dilemma, since I reserve 'wrong' for the strongest possible criticism in the relevant kind of situation" (1988: 20). Elsewhere he draws attention to the fact that, although the term "wrong" is usually reserved for acts which violate overriding moral requirements, at times it is also used to refer to violations of non-overridden moral requirements (1996: 53–4). Sinnott-Armstrong insists that moral dilemmas are to be conceived of as conflicts between non-overridden moral requirements (1988: 17–8). If one adopts this definition of dilemmas and the former usage of the term "wrong", what follows is that acting in a moral dilemma involves no moral wrongdoing. However, some proponents of moral dilemmas use the term "wrong" differently. For instance, Raz holds

<sup>&</sup>lt;sup>11</sup> For emphasizing this point, see Foot (2002).

<sup>&</sup>lt;sup>12</sup> It is worth stressing that proponents of dilemmas accept that some moral conflicts leave no moral residue. See, for instance, Sinnott-Armstrong (1996: 54–5) and Railton (1996: 153–4). To use Sinnott-Armstrong's example, suppose X promises to meet Y for lunch, but breaks his promise in order to save someone's life. X finds out later that, if he had met Y at the restaurant, they would have both been killed by a bomb. In this case, it would be odd to say that X has an obligation to apologize or to compensate Y. Yet, opponents of dilemmas insist that moral conflicts are *always* soluble without a remainder.

<sup>&</sup>lt;sup>13</sup> For more on this point, see Gowans (1994: 96).

that there are cases when the right thing to do involves some form of wrongdoing, and that it would be a mistake to believe that performing what is under the circumstances the right action erases its wrongful character (1986: 359, 365). Clearly, those who embrace this position do not restrict the use of the term "wrong" to the strongest possible moral criticism. Despite insisting that acting in a moral dilemma involves doing a wrong, the term "wrong" is not used here in the sense of acting against an all-things-considered moral ought.

Third, as long as we do not use the term "wrong" in the sense of failing to act in accordance with the correct conclusion of moral deliberation, we can say that proponents of dilemmas agree that acting in a moral dilemma involves doing something wrong. The idea that, even by acting in accordance with the correct conclusion of moral deliberation, one may do something which is in some sense morally wrong is captured by the claim that the resolution of moral conflicts leaves a moral residue. By contrast, opponents of dilemmas deny that there are cases when acting in accordance with the correct conclusion of moral deliberation involves doing something which is, in any sense, morally wrong. Hence, their insistence on the fact that the resolution of moral conflicts leaves no moral residue.<sup>14</sup> The position advocated by friends of dilemmas lies between two extremes. At one extreme, we find the position advocated by foes of dilemmas, who claim that insofar as the agent respects the correct conclusion of moral deliberation (i.e. insofar as the agent acts in accordance with a correct *all-things-considered* moral judgment), it is not possible for him to do something which is morally wrong in any sense. So, foes of dilemmas reject the possibility of inescapable moral wrongdoing. At the other extreme, we find a position that is often erroneously attributed to friends of dilemmas, according to which there are cases when one cannot escape wrongdoing in the sense of violating the correct conclusion of moral deliberation. While friends of dilemmas contend that there are cases when moral wrongdoing is unavoidable, this claim should not be taken to imply that one cannot avoid violating an all-things-considered moral ought.

Yet, one can draw attention to the fact that there are significant differences between the views proposed by different advocates of moral dilemmas and object that the present approach levels out these differences. All advocates of moral dilemmas endorse the view that the resolution of moral conflicts leaves a moral remainder. However, some of them argue that the remainder thesis is uninteresting insofar as it concerns residual obligations, and seem unwilling to go further and endorse the claim that there are situations when one cannot escape wrongdoing.<sup>15</sup> By contrast, other advocates of moral dilemmas insist that moral wrongdoing is sometimes inescapable.<sup>16</sup> Still, I believe that these differences are at least partly explained by the fact that "wrong" is used differently and that clarifying the meaning of this term is the best way to shed light on this debate. Moreover, endorsing the reminder is thesis is enough to yield the conclusion that there are cases when one cannot escape wrongdoing (in the sense discussed above). It is also worth emphasizing that even those who maintain that the remainder thesis is uninteresting will have to admit that one might be confronted with cases in which committing a grave moral wrong is inescapable. (Again, committing a grave moral wrong should not be equated here with violating an all-things-considered moral ought.) This conclusion follows naturally if the remainder

<sup>&</sup>lt;sup>14</sup> For a similar characterization of the debate about moral dilemmas, see Gowans (1994: 90–1). <sup>15</sup> See, for instance, Foot (1983: 396; 2002).

<sup>&</sup>lt;sup>16</sup> See, for instance, Raz (1986).

thesis is accepted. If one accepts that conflicting *prima facie* obligations make genuine moral claims, then one should also accept that the resolution of moral conflicts can leave a great moral remainder, since how great the moral remainder is simply depends on the weight of the unfulfilled *prima facie* obligation. As János Kis points out:

The harm that S may cause to other people when she fails to satisfy a relatively trivial ought is likely to be a relatively trivial harm. Relatively trivial harms are likely to be relatively easy to repair. The less dramatic the ought, the more likely it is that the victims can be fully compensated. And conversely: the more dramatic the ought, the more likely it is that the victims cannot be fully or even partially compensated. (Kis, 2008: 255)

I have already emphasized that my aim here is not to provide a definition of moral dilemmas, but merely to identify several claims that are endorsed by most proponents of dilemmas. Those who defend the existence of moral dilemmas argue that there are cases in which there are no moral grounds for choosing between two incompatible courses of action, each of which is supported by some moral reasons. Furthermore, proponents of dilemmas share the view that the resolution of moral conflicts leaves a moral remainder.<sup>17</sup> On their view, it would be a mistake to think that, as long as the agent acts in accordance with the correct conclusion of moral deliberation (i.e. in accordance with a correct allthings-considered moral judgment), it is not possible for him to do something which is morally wrong in any sense. The idea conveyed by remainder thesis is that, even by acting in accordance with the correct conclusion of moral deliberation, the agent can do something which is in some sense morally wrong. Therefore, as long as we do not use the term "wrongdoing" in the sense of acting against an all-things-considered moral ought, we can attribute to proponents of dilemmas the view that moral wrongdoing is sometimes unavoidable.<sup>18</sup> Moreover, as long as "wrongdoing" is not used in the above-mentioned sense. I think it is safe to say that most proponents of dilemmas would admit that there are cases in which committing a grave moral wrong is inescapable. If the remainder thesis is accepted, there seems to be nothing to bar this conclusion. By contrast,

<sup>&</sup>lt;sup>17</sup> One reason against defining moral dilemmas by appeal to the notion of moral residue is that in a number of cases that we see as paradigmatically dilemmatic the only person who is harmed is the agent himself. In such cases, it does not make sense to bring into discussion obligations to explain, apologize or compensate. For stressing this point, see Foot (2002: 183) and Kis (2008: 257). However, finding a definition of moral dilemmas that would cover all cases that we think of as genuinely dilemmatic is a notoriously difficult enterprise, in which I do not engage here.

<sup>&</sup>lt;sup>18</sup> Some might raise doubts about whether the claim that an overridden obligation is a real obligation which persists as a moral remainder should be taken to imply that the failure to discharge such an obligation is in some sense morally wrong. On this reading, every moral conflict would involve some form of wrongdoing. So the question is whether advocates of dilemmas would want to endorse such a strong view. Three points are worth emphasizing here. First, the remainder thesis is intended to apply to conflicts of obligations. Thus, in order to claim that failing to perform an act is in some sense morally wrong, it has to be shown first that one has a moral *obligation* to perform the act in question. Second, proponents of dilemmas stress that not all conflicts of obligations leave a moral remainder. See also *supra* note 12. Finally, as we have already seen, advocates of dilemmas maintain that there is a difference between violating a *pro tanto* moral ought and violating an all-things-considered moral ought. (Whether the terminology they employ is appropriate is a different issue.) Also, advocates of dilemmas accept that conflicts of obligations can leave a greater or a lesser moral remainder. In short, the view considered here is less extravagant than it may seem at first glance.

opponents of dilemmas maintain that, barring certain exceptional cases, there are no irresolvable moral conflicts. Also, they insist that moral conflicts are always soluble without a remainder. Given that, on this view, *prima facie* obligations make no genuine moral claims, there are no grounds for holding that failing to discharge such an obligation can be morally wrong in any sense.

In what follows, I will make a few brief remarks on the relationship between value incommensurability and moral dilemmas. As I have already pointed out, I will not take a stand on whether we should reserve the term "moral dilemma" for irresolvable moral conflicts or use it instead to refer to cases in which one cannot escape committing a grave moral wrong. I will argue that, regardless of how we choose to define moral dilemmas, incommensurability still plays an important role in accounting for them.

## 5.4 Value incommensurability, irresolvable moral conflicts and moral wrongdoing

A good deal of confusion in the literature on moral dilemmas is due to the failure to distinguish among two different arguments that relate incommensurability to dilemmas. Those who claim that incommensurability establishes the possibility of moral dilemmas do not always make it clear whether incommensurability is intended as an account of irresolvable of moral conflicts, or rather as an account of inescapable moral wrongdoing. Most of their critics argue that incommensurability does no essential work in accounting for dilemmas precisely because it leaves unexplained a feature that they see as crucial for moral dilemmas, namely there being no course of action entirely free from wrongdoing. By considering these two arguments, I will try in what follows to cast light on this debate.

I turn now to the first argument, according to which value incommensurability accounts for the possibility of irresolvable moral conflicts.<sup>19</sup> This argument can be stated as follows: if the value of two conflicting options cannot be compared in a way that is pertinent to moral deliberation, then there are no moral grounds for choosing one option over the other. Yet, one can draw attention to the fact that, while this is a valid argument, it is not a particularly informative argument. It seems, however, that those who talk about incommensurability in the context of dilemmas intend to say something substantive about why moral dilemmas arise. It is worth stressing, for instance, that some of these authors give an account of incommensurability in terms of abstract values, rather than in terms of concrete options.<sup>20</sup> Note also that what these authors usually claim is that dilemmas arise because values are plural and incommensurable. In other words, the claim that values are irreducibly plural plays a key role in the argument.

However, it is important to emphasize that several commonly accepted versions of this argument are misguided. Those who hold that moral dilemmas arise because values are plural and incommensurable often have in mind Isaiah Berlin's idea that there is an inherent conflict between values.<sup>21</sup> According to Berlin, not all the values that are worth

<sup>&</sup>lt;sup>19</sup> See, for instance, Nagel (1979: 128–9), Foot (1983: 395–7), Raz (1986: 357–66), and Sinnott-Armstrong (1988: 58–71).

<sup>&</sup>lt;sup>20</sup> See, for instance, Nagel (1979).

<sup>&</sup>lt;sup>21</sup> For claiming that the view according to which all enduring values can be harmoniously pursued is "conceptually incoherent", see Berlin (1959: 15).

pursuing can be realized in a single life or a single society. This is because securing one value often involves frustrating the pursuit of another value. For instance, one cannot at the same time be both a mother and a nun, nor can one society be committed both to the redistribution of wealth and to economic *laissez faire*. It is precisely because we are confronted with choices among such rival goods, the argument goes, that moral dilemmas arise.<sup>22</sup> Yet, insofar as it proceeds from the claim that values are incompatible, the argument must be rejected. There are two reasons why this is so.

First, the incompatibility of values does not amount to their incommensurability.<sup>23</sup> Presumably, one cannot become as skilled a smith as one is a neurosurgeon, since continuing to be the latter requires taking care of one's hands in a fashion that is incompatible with becoming the former. Still, from the fact that being a neurosurgeon precludes the option of being a smith, it does not follow that no relationship as to which one is more valuable holds between the two.

A brief digression into the debate on moral dilemmas will reveal a second, more important reason why the tenet that there is an *inherent* conflict between values cannot account for the possibility of moral dilemmas. On the one hand, opponents of dilemmas have argued that the existence of dilemmas is evidence of inconsistency among the moral principles that give rise to them. For instance, Donagan maintains that "the generation of moral dilemmas is to moral rationalism what the generation of self-contradictions is to theories generally: an indispensable sign that a particular theory is defective" (1996: 15). On the other hand, proponents of dilemmas attempt to define inconsistency in such a way as to prove that dilemma-generating theories are not lacking in this particular respect. Notably, Ruth Marcus proposes to "define a set of rules as consistent if there is some possible world in which they are all obevable in all circumstances in *that* world" (1980: 128). Hence, "a set of rules is inconsistent if there are no circumstances, no possible world, in which all the rules are satisfiable" (1980: 129, all emphases in original). On this interpretation of consistency for a set of moral principles, the possibility of dilemmas is not indicative of a defective moral view, since moral dilemmas might simply arise due to *contingent* factors.

Even though proponents of dilemmas have initially confined the debate to cases in which moral conflict has a contingent basis<sup>24</sup>, subsequent discussions have focused more and more on the idea that there is an inherent conflict between values. I suggest that this shift was ill-advised. As long as proponents of dilemmas insist that moral theories which do not provide for the resolution of every possible *contingent* conflict are not defective, it is not open to them to argue that dilemmas arise due to intrinsic conflicts between values. If there is an intrinsic conflict between two values, in the sense that realizing one of them would necessarily lead to failing to realize the other, then there is no conceivable world in which anyone could realize both of them. Therefore, identifying this type of conflict as the source of moral dilemmas does not help the proponents of dilemmas, since it points to what is, by their own lights, an inconsistency.

Yet, one can try to restate the case as follows. An adequate understanding of human values comes with the realization that different forms of life display *genuine*, though *incompatible* virtues. While admitting, for instance, that one cannot at the same time be

<sup>&</sup>lt;sup>22</sup> For this suggestion, see Foot (1983: 396–7).

 $<sup>^{23}</sup>$  For this distinction, see Williams (1981: 76–7), Kekes (1993: 22), and Griffin (1997: 36).

<sup>&</sup>lt;sup>24</sup> See also Williams (1973*a*: 171).

both a mother and a nun, we acknowledge that these forms of life display distinct virtues, each of them being worth pursuing for its own sake.<sup>25</sup> Taking this stance does not exhibit any logical inconsistency. It is only that, on reasonable assumptions about human nature, one cannot pursue both callings. Thus, one can argue that the most defensible normative outlook is one that endorses the existence of a wider range of human values than those which can be realized in the life of a single individual. And one might suggest that the occurrence of moral dilemmas can be traced back to there being such a generous range of values genuinely worthy of pursuit.

It is important to stress, however, that opponents of dilemmas do not find puzzling the idea that a moral predicament can take the form of a moral dilemma when one is faced with conflicting requirements that originate from independent sources of obligation. As Dworkin (2001) puts it, as long as one sees oneself as the subject of "two independent sovereigns", it is not surprising that one will be confronted with difficult moral choices.<sup>26</sup> To counter Dworkin's suggestion that moral dilemmas can be explained away, friends of dilemmas have to show that a narrow range of values can give rise to genuine dilemmas.

Another common version of the argument that relates incommensurability to moral dilemmas elicits the claim that values are incommensurable from the claim that they are irreducible to one another. For instance, Thomas Nagel (1979) argues that the occurrence of moral dilemmas can be traced back to there being distinct types of moral values which are incommensurable. He holds that there are "personal" or "agent-centered" values, i.e. values that derive from obligations, rights and commitments to one's own projects, and "impersonal" or "outcome-centered" values, i.e. utility and perfectionist ends, and that it is the fact that these five fundamental types of values are irreducible to a common basis that explains why conflicts that arise between them are irresolvable. If this argument fails, it fails for the same reason why arguments for incommensurability that rely on the claim that values are irreducibly plural seem to fail. Simply put, the idea is that some options that instantiate distinct types of values are comparable and that arguments from pluralism "are not sufficiently fined-grained to differentiate cases of putative

<sup>&</sup>lt;sup>25</sup> This example comes from Raz (1986: 395).

<sup>&</sup>lt;sup>26</sup> Dworkin (2001: 81-2, 138-9) cites the example of Abraham and Isaac and argues that what explains Abraham's conflict is the fact that God and morality are viewed as "two independent sovereigns". In this context, it is also worth emphasizing that the disagreement between Dworkin and proponents of dilemmas goes deeper than a difference of opinion about whether one can be faced with morally incommensurable alternatives. Dworkin contends that, unless it is proved that the most attractive conceptions of two given values produces a conflict, it begs the question to say that the values in question conflict. (Dworkin has repeatedly emphasized, for instance, that the point that should not be conceded to value pluralists is not the lack of a further perspective from which to reconcile two fundamental values which conflict, but the more basic point that such values, properly understood, can conflict. See Dworkin (1996c; 2001). For a similar view, see White (2004).) For criticizing Dworkin's view that the appearance of conflict between two values can always be explained away by a proper interpretation of the values in question, see Nagel (2001) and Williams (2001). Nagel (2001: 129) stresses, for instance, that in the case of the debate over the allocation of public resources between social justice and scientific research, it would be absurd to maintain that a redefinition of the value of scientific research would eliminate the conflict. Moreover, even when it is possible to redefine the relevant values, it is doubtful whether such attempts can succeed in making values immune to conflict. As Williams argues, "the concerns which basically go with these various values cannot be redirected simply nominalistically, by redefining a word" (2001: 94).

incomparability from ones of certain comparability" (Chang, 1997: 15).<sup>27</sup> While arguments from pluralism have been widely criticized, most authors agree that incommensurability requires pluralism in the theory of value.<sup>28</sup> Therefore, attempting to come up with a version of the argument that avoids the difficulty mentioned above would constitute a valuable philosophical exercise. Yet, even if no such argument is available at the moment, we have already seen in Chapter 4 that there is a convincing argument to the effect that the value of certain options cannot be ranked against each other as "better than" or "equally good". So let me focus now on the relationship between incommensurability and moral dilemmas.

As it has been frequently emphasized, incommensurability is neither necessary nor sufficient for moral dilemmas to exist. Moral dilemmas can arise when the moral reasons in favor of two incompatible courses of action are evenly balanced. Moreover, one can point out that the moral reasons on each side might not be sufficiently strong to yield an obligation.<sup>29</sup> However, I believe that incommensurability still plays a role in accounting for moral dilemmas. Let me explain. In order to establish that irresolvable moral conflicts are possible, proponents of dilemmas often emphasize that even a single moral principle can generate dilemmas.<sup>30</sup> For example, one can promise two colleagues to help with their workload and then discover that circumstances beyond one's control make it impossible to keep both promises. Admittedly, whenever there is a moral conflict that is perfectly symmetrical, there is no moral ground for choosing one option over the other. Opponents of dilemmas allow for the possibility that the moral reasons in favor of two incompatible courses of action are equally compelling.<sup>31</sup> Nonetheless, they typically insist that cases in which competing moral reasons are evenly balanced are rather exceptional. Furthermore, they reject the idea that moral reasons can be incommensurable. The upshot of all these considerations is that both parties to the debate admit that the frequency of irresolvable moral conflicts will depend on whether the idea of incommensurability is accepted.<sup>32</sup> As we have already seen, opponents of dilemmas do not dispute the theoretical possibility of moral ties. However, given that proponents of dilemmas maintain that moral conflicts can be irresolvable in cases when the conflict is not symmetrical, incommensurability would play an important role in accounting for moral dilemmas.

<sup>&</sup>lt;sup>27</sup> See also Regan (1997: 135). For criticizing Nagel's version of the argument, see also Sinnott-Armstrong (1988: 60–1) and Dancy (1993: 261–3).
<sup>28</sup> A notable exception is Chang (1997: 16–7). For claiming that the issue of monism and pluralism cuts

<sup>&</sup>lt;sup>28</sup> A notable exception is Chang (1997: 16–7). For claiming that the issue of monism and pluralism cuts across the issue of commensurability and incommensurability, see also Griffin (1986: 89–92; 1997: 36).

<sup>&</sup>lt;sup>29</sup> For this point, see Conee (1989: 139–140). Consider, for instance, the dilemma faced by the student in Sartre's example. It is not at all clear that the most compelling interpretation of the student's plight is couched in terms of conflicting moral *obligations*. Admittedly, a moral conception that claimed that every young man in France had the moral obligation to join the military would be overdemanding. Even if one accepts a general duty to contribute to the defeat of the Nazis, one can still envisage other ways in which the student could have discharged this duty. Surely, there were other, innumerable ways of helping the cause while staying in France. Moreover, one can argue that staying by one's aging parent's side whatever the circumstances is not an act of filial duty, but rather a supererogatory act. For this interpretation of the dilemma presented by Sartre, see Railton (1996: 149–151). Railton acknowledges that some cases of moral conflict that we see as paradigmatically dilemmatic do not involve a clear conflict of obligations. However, he argues that this is rather a reason for broadening our understanding of dilemma to include conflicting moral ideals.

<sup>&</sup>lt;sup>30</sup> See, for instance, Sinnott-Armstrong (1988: 54–7) and Marcus (1996: 25).

<sup>&</sup>lt;sup>31</sup> See, for instance, Donagan (1984: 307–8).

<sup>&</sup>lt;sup>32</sup> For this point, see also Brink (1996: 107).

Yet, one can grant that the frequency of irresolvable moral conflicts depends on the existence of moral incommensurables, and still argue that value incommensurability has nothing essentially to do with moral dilemmas. It is worth stressing that this objection can come from both opponents and proponents of dilemmas. Some proponents of dilemmas do not see irresolvability as a necessary condition for moral dilemmas.<sup>33</sup> On their view, dilemmas can occur even when the moral reasons in favor of one course of action clearly override the moral reasons in favor of the other available course of action. What seems to be required for the occurrence of dilemmas is rather that there is a great moral residue. However, one can contend that incommensurability plays no role in accounting for moral residues or inescapable moral wrongdoing. And one can insist that it is precisely the idea of inescapable moral dilemmas allow for the possibility of irresolvable moral conflicts, but reject the possibility of inescapable moral wrongdoing.

Although at first it may seem that incommensurability has nothing to do with the possibility of unavoidable moral wrongdoing, the former can help explaining the latter. Here is a suggestion that comes from Raz (1986). He argues that:

It is question-begging to assume that a person can only do wrong by failing to perform a better act which he could have done. Anyone asserting the existence of dilemmas [in which there is a right course of action] is denying precisely that. The claim that there are such dilemmas is a claim that one can do wrong in other situations as well. Incommensurability shows that there is conceptual room for a notion of wrongdoing which does not involve failure to take a better action available to one. (Raz, 1986: 360)

Raz goes on to emphasize that those who insist that in every situation there is at least one option that is entirely free from wrongdoing do not judge the value of options, but give advice about the eligibility of an option compared to the others. However, in cases when the agent is confronted with a choice between options that are incommensurable, the aforementioned kind of judgment is ruled out. As Raz puts it, in such cases "one cannot compare the value of the options, one can only judge their value each one on its own" (1986: 364). If each option involves some form of wrongdoing, then one cannot avoid the conclusion that wrongdoing is inescapable. Thus, incommensurability proves that there is conceptual room for holding that one can do a wrong without a failure to perform a better action which was available. In short, Raz shows that incommensurability plays a key role in answering the objection according to which "saying of an action one cannot escape that it is wrongful is idle and confusing talk" (1986: 365).

### **5.5 Conclusion**

My aim in this dissertation has been to provide an answer to the question whether moral objectivists can adequately respond to the argument from disagreement. According to this argument, the pervasiveness and intractability of moral disagreement gives us reason to

<sup>&</sup>lt;sup>33</sup> See Williams (1973*a*), Raz (1986), Marcus (1996), and Foot (2002).

reject moral objectivism. A common suggestion in the recent metaethical literature is that the argument from disagreement loses its bite once we concede that some moral issues are indeterminate. The bulk of this dissertation has been devoted to answering the question whether the appeal to the idea of moral indeterminacy can help the objectivist win this debate.

I have argued that this response to the argument from disagreement is not free of difficulties. I have pointed out that, since moral indeterminacy cannot be a pervasive phenomenon, the best strategy available to the objectivist is to maintain that there is also a considerable amount of convergence in moral belief, and that this convergence in belief is best explained on objectivist assumptions. Yet, if responding to the argument from disagreement involves defending both the claim that there are indeterminate moral issues and the claim that convergence in moral belief can be regarded as the result of a reliable epistemic mechanism, a worry might arise as to whether an adequate defense of both the aforementioned claims can be mounted from the perspective of the same version of moral objectivism. The worry is that defending the former claim may be easier if one embraces a version of cognitivism which delivers a weaker form of objectivity, while defending the latter claim may be easier if a version of cognitivism which delivers a stronger form of objectivity is vindicated. My aim throughout this dissertation has been to show that an adequate defense of both the aforementioned claims can be mounted claims can be mounted from the perspective of moral realism.

On the one hand, I have argued that, despite common assumptions to the contrary, moral realism has no trouble accommodating moral indeterminacy. More specifically, I have distinguished between comparative and alethic indeterminacy and argued that neither of them poses a threat to moral realism. On the other hand, I have argued that it would be wrong to think that moral facts are causally inert, and therefore, they cannot figure in a causal explanation of our moral beliefs. Moreover, I have shown that the claim that moral facts can play causal explanatory roles can consistently be endorsed by both naturalists and nonnaturalists.

Furthermore, I have pointed out that those who maintain that the argument from disagreement can be refuted by appealing to the idea of moral indeterminacy seem to face an insuperable difficulty, namely that absent a belief in there being a determinate answer to the question under consideration, it does not make sense to disagree. In other words, one can stress that agents who suffer from no cognitive shortcomings (i.e. agents who are aware of all morally relevant facts, including those which can render a situation morally indeterminate) would have no reason to disagree. Yet, moral realists can insist that the crux of the matter is not whether there can be disagreement among agents who suffer from no cognitive shortcomings, but whether such agents can sometimes fail to converge on a single moral verdict. I have argued that shifting the focus from whether there can be disagreement among ideal agents to whether such agents can sometimes fail to converge on a single moral verdict should be allowed to count as a legitimate move in responding to the argument from disagreement.

Finally, I have emphasized that the realist who attempts to respond to the argument from disagreement cannot avoid engaging in certain first-order ethical debates. I have pointed out, for instance, that, insofar as the plausibility of the realist's response to the argument from disagreement depends on whether there are cases of comparative moral indeterminacy, the realist cannot avoid dealing with the question whether there are cases when the traditional trichotomy of relations "better than", "worse than" and "equally good" fails. Although the view according to which there are cases when none of the traditional value relations obtains is highly controversial, there is strong reason to believe that incommensurabilists have the upper hand in this debate. Therefore, the realist who claims that there are indeterminate moral issues can derive support from arguments put forward by incommensurabilists. I have also argued that, in order to dispel the worry that the view according to which some options are morally incommensurable is merely the result of transplanting a justification designed for one domain (e.g. the artistic domain) to a different domain (i.e. the moral domain), the realist can appeal to considerations about moral dilemmas. A close look at the debate on moral dilemmas shows that the suggestion according to which the view that there are cases of comparative indeterminacy reflects a neglect of substantive moral issues is misleading.

To sum up, there is at least one version of moral objectivism, i.e. moral realism, that has the resources to answer the argument from moral disagreement. Yet, in order to offer a convincing reply to this argument, the moral realist cannot avoid addressing other metaethical and ethical issues as well.

# Appendix Open texture

### 1 Waismann's notion of open texture

The phenomenon of open texture, allegedly a pervasive feature of language, "reflects the fact that we cannot take into account ahead of time all the possible situations in which our words might be put into test" (Margalit, 1979: 142). Simply put, the idea is that since we cannot envisage all possible circumstances in which our terms may be applied, we cannot come up with exhaustive definitions of our terms. As a result, most of our terms are open textured, i.e. in unanticipated circumstances, their application can become indeterminate.

Originally, the notion of open texture was invoked by Friedrich Waismann (1945) in connection with empirical terms.<sup>1</sup> His argument was aimed at refuting verificationism. Waismann pointed out that, since we cannot foresee all possible circumstances in which our empirical terms can be applied, we cannot specify all possible circumstances in which sentences containing such terms are true or false. He concluded that empirical sentences cannot be conclusively verified and that the verifiability criterion of meaningfulness should therefore be rejected.

According to Waismann (1945: 123), open texture should be distinguished from the related linguistic phenomenon of vagueness. A term is vague if and only if it allows for borderline cases, i.e. cases in which the question whether the term applies or fails to apply has no determinate answer. Terms like "bald", "heap" or "middle-aged" typically allow for such borderline cases. Open texture is not currently identifiable vagueness, but rather *potential vagueness*, a property that Waismann attributes to all empirical terms. He claims that, in unanticipated circumstances, any term's application can become indeterminate. Supposing we came across a substance which had the chemical composition of gold, but emitted a new sort of radiation, or a catlike creature which could be revived from death, we would be uncertain whether the term "gold" or "cat" can legitimately be applied in such cases.

There is, however, another characterization of the distinction between open texture and vagueness, which is not explicit in Waismann's writings, but which may help clarify the exposition of his ideas. As already pointed out in Chapter 2, there seem to be at least two distinct types of vagueness: "sorites vagueness" or "degree vagueness", which stems from the lack of a sharp cutoff point along some continuum, and "combination of conditions vagueness", which occurs when there is no determinate answer to the question of what combination of conditions is necessary or sufficient for the application of a term.<sup>2</sup>

<sup>&</sup>lt;sup>1</sup> Later, the notion of open texture was imported into legal theory. H. L. A. Hart (1961: ch. 7) used the thesis that all language is open textured as an argument for judicial discretion. He claimed that legal rules embody choices made with respect to foreseeable circumstances. However, he emphasized that, when legal rules are to be applied to unenvisaged circumstances, a further exercise of choice is needed. For an interesting discussion of open texture as applied to moral concepts, see Brennan (1977).

 $<sup>^2</sup>$  Traditionally, the analysis of vagueness was associated with the search for a solution to the sorites paradoxes (e.g. one grain of sand does not form a heap; moreover, adding just one more grain of sand to something that is not a heap is insufficient to form a heap; hence, no matter how large the number of grains of sand, it will not create a heap). Vagueness was analyzed in terms of an underlying scale along which

While degree vagueness occurs in terms like "bald", "tall" or "middle-aged", "religion" or "sexual perversion" are examples of terms that exhibit combination of conditions vagueness. In claiming that most of our empirical terms are potentially vague, Waismann had in mind the latter type of vagueness.<sup>3</sup> As it will soon become clear, the phenomenon of open texture can be traced back to our inability to anticipate all possible combinations of conditions in which a term can be applied.

Waismann (1945: 123) claims that open texture cannot possibly be eliminated from our linguistic apparatus. More concretely, he argues that *potential* vagueness, as opposed to *actual* vagueness, cannot be remedied by giving more accurate rules. The explanation he offers clearly suggests that, in discussing the phenomenon of open texture, he has in mind combination of conditions vagueness. Waismann points out that we cannot construct *complete definitions* of our empirical terms, which would anticipate and settle of an ultimate manner all possible questions concerning their future usage. As he rightly emphasizes, when we define an empirical term, we have in mind only certain kinds of circumstances. So, what we do is to render the term precise along some dimensions of its meaning. However, not all possible circumstances in which a term is to be used can be foreseen. The possibility that we have not taken into account something that is relevant to the term's usage, something that would make us extend or modify its previous definition, is always present. It is along these dimensions which have not been considered that the application of the term might become indeterminate.

In Waismann's view, two factors combine to create the open texture of our empirical concepts: *the incompleteness of empirical descriptions* and *the incompleteness of our factual knowledge* (1945: 124–7). Part of the reason for our not being able to come up with an exhaustive list of the circumstances in which a term can be applied is that the description of a material object or of a situation can never be exhausted: no matter how carefully detailed a description is, one can always add something further to it. Yet, Waismann argues that it is our limited knowledge of the possible circumstances in which we would need to *revise* or *retract* particular statements employing an empirical concept that has the most relevant stake in accounting for the phenomenon of open texture. He rightly points out that if the possibility of something new and unforeseen emerging were not present, the empirical description of the situations in which a term is to be used would be incomplete only in the sense that it could never be finished.

differences between adjacent points are imperceptible. However, there are terms having borderline cases of application for which there is no such scale. For instance, there might be no determinate answer to the question whether a certain social practice constitutes a religion despite the fact that discriminating how many of the relevant features are displayed by the practice in question involves no difficulty. This has led theorists to distinguish between "sorites vagueness" or "degree vagueness" and "combination of conditions vagueness". See, for instance, Alston (1964: 87–8).

<sup>&</sup>lt;sup>3</sup> Some commentators argue that, in fact, the philosophical import of Waismann's notion of open texture collapses into the philosophical import of combination of conditions vagueness. See, for instance, Ackerman (1994: 130). Ackerman claims that her interpretation does not disparage Waismann's contribution in showing that even terms that do not actually give rise to borderline cases are potentially vague.

### 2 Open texture and rule-following

Waismann's point about what constitutes a challenge to the established use of a concept is spelled out by Avishai Margalit, by means of a distinction between *open texture* and *open-endedness*. Margalit (1979: 142) points out that we cannot possibly foresee all the circumstances in which a term such as "intelligent" can be applied. As he emphasizes, a dispositional term such as "intelligent" seems to be more *open-ended* than a term such as "alcoholist", i.e. while there are many different types of behavior that an intelligent person can display, there is one type of behavior that can be expected from an alcoholist. Still, the fact that an individual manifests a behavior which no one has previously thought of as "intelligent" need not hinder us from classifying it as such. Doubts concerning the correct application of the term appear only to the extent that some of our basic beliefs about the behavior of an "intelligent" person are violated. Thus, *open texture* should be reserved "to cover those cases in which some of our hard core beliefs are assumed to be violated" (Margalit, 1979: 143).

Margalit's interpretation seems to be supported by Waismann's examples. For instance, if we come across a catlike creature that could be revived from death, shall we say that a new species has come into being, or rather that we are confronted with a cat with extraordinary properties? Our irremediable puzzlement concerning the application of a natural kind term to an entity which possesses most of the features shared by other members of the group, but which also possesses features that disqualify it from belonging to the group, stems from the fact that our hard core beliefs about the nature of the kind in question seem to be violated. However, this interpretation of Waismann's notion of open texture might be problematic, given the difficulties inherent in attempting to demarcate hard core beliefs from non-hard core ones and the difficulties involved in attempting to specify the sense in which a Wittgensteinian scholar like Waismann could speak about "hard core beliefs". Moreover, it seems reasonable to suppose that Waismann would have endorsed the claim that indeterminacy of the open texture variety arises also in cases that involve no violation of our beliefs. An example will serve to drive this point home. Consider the term "velocipede", which denominates a forerunner of the modern bicycle. The velocipede was constructed of wood and iron, and had a large front wheel and a small rear wheel. When a similar vehicle with two wheels of an equal diameter was invented, people could have applied the term "velocipede" to this new machine, or they could have devised a different term to refer to it. The previous definition of the term left this aspect simply uncovered. As Waismann would put it, when we defined the term "velocipede", we framed its application conditions with respect to foreseeable situations, and so we did not consider the question whether the term can properly be applied to a vehicle with two wheels of an equal diameter. Hence, the indeterminacy of application. It should be noticed, however, that the invention of what came to be known as the bicycle did not pose a challenge to any of the beliefs held at the time.

Thus, it seems we should concede that in discussing open texture Waismann had in mind a range of cases that included both circumstances in which some of our most deeply entrenched beliefs would be violated and circumstances in which no such violation is involved. Still, focusing primarily on the later kind of circumstances might obscure what is distinctive about Waismann's notion of open texture. We might be misled into thinking that Waismann's considerations about open texture constitute a mere rehearsal of Ludwig Wittgenstein's considerations about rule-following, and fail to notice that what makes Waismann's notion of open texture distinctive is precisely the claim that the potential vagueness of our terms is rooted in the incompleteness of our empirical knowledge. Let me detail.

In his *Philosophical Investigations* (§ 185ff), Wittgenstein presents us with the case of a pupil who is instructed to extend the mathematical series 0, 2, 4, 6, 8, ... by adding 2 each time. After reaching 1000, the pupil continues the series with 1004, 1008, 1012, .... Our natural reaction to this example is to think that the pupil did not understand the teacher's instructions, since the rule "add 2" determines a unique sequence of numbers which is to be carried on. Still, Wittgenstein asks us to imagine that to any attempt to clear up the misunderstanding the pupil will reply that he did what he was asked to do. For instance, he could reply that he understood the instruction "add 2" as "add 2 if the resultant sum is less than 1000, but add 4 if the resultant sum is greater than 1000". The point is that a rule does not determine its correct application. Since any formulation of a rule is open to different interpretations, different courses of action can be made out to accord with the rule. Just as in the case of extending a series of numbers what is to "carry on in the same way" is not uniquely determined, the past applications of a term (or the rule governing its application that can be formulated starting from them) do not uniquely determine its future applications.

Wittgenstein's considerations about rule-following are aimed at rebutting the view that understanding a term's meaning consists in grasping a rule which, in turn, guides our application of the term in an infinite number of cases. The exploration of cases such as the one mentioned above, in which the pupil's reaction to his training is "abnormal", is intended to highlight certain aspects of the normal cases which we tend to misunderstand. This example shows that a rule does not compel a particular application of it, since there are different things that we might call applications of the rule. It should be emphasized, however, that even if Wittgenstein is right in suggesting that there are no rules that would uniquely and unequivocally determine the application of our terms, it does not follow that our linguistic behavior is completely unconstrained. As it has often been pointed out in the literature, reading Wittgenstein as a skeptic about rules and meaning seems to misrepresent his thoughts.<sup>4</sup> On a more plausible reading, Wittgenstein's aim is to free us from the temptation to think that it only makes sense to speak of a "correct" use of our concepts if the way they are to be used is uniquely predetermined. In choosing a case in which there is a consensus about the correct way to go on, he is interested in pinpointing what it is exactly that makes the continuation of a mathematical series or the application of a term correct. According to him, the distinction between correct and incorrect use is grounded in our practice of employing a rule, and not in the rule itself.

If we focus on Wittgenstein's idea that the meaning of a term does not anticipate all its future applications, we might think that Waismann's considerations about open texture convey the same idea. There are obvious similarities between the two approaches. On the one hand, it makes no sense to claim that our terms are open textured unless we assume that the Platonic imagery of "rules as rails" is seriously inadequate. In his *Philosophical Investigations* (§ 218–9), Wittgenstein argues that even though we are naturally inclined

<sup>&</sup>lt;sup>4</sup> For a skeptical interpretation of Wittgenstein's considerations about rule-following, see Kripke (1982). For a different interpretation, see Baker and Hacker (1984), McDowell (1998*a*; 1998*b*), and McGinn, M. (1997).

to think of rules as rails stretching ahead of us to infinity, this picture misrepresents the nature of following a rule. The picture of rails is used in this context to express the idea that "all the steps are really already taken" (§ 219). Clearly, this Platonic picture of rules is inconsistent with claiming that our terms are open textured. The point is not only that the above-mentioned picture implies that at each new step there is a single correct way of applying a rule (since all the steps are already taken), but also that we are supposed to let ourselves be guided by these rails stretching invisibly to infinity. Thus, if understanding a term's meaning consisted in grasping a rule that would guide our application of the term in an infinite number of cases, then the mere fact that the definition of a term did not incorporate everything that might prove to be relevant to its usage would not constitute an obstacle to our knowing how to go on in applying it. On the other hand, Wittgenstein also insists in the *Investigations* that "it is only in normal cases that the use of a world is clearly prescribed" (§ 142). He adds that the more unusual the context, the more doubtful it becomes whether the term applies or fails to apply.

It should be emphasized, however, that Waismann's view about open texture differs from Wittgenstein's view about rule-following in two important respects. In pointing out that a rule does not compel a particular application of it, Wittgenstein's interest lies mainly with *familiar*, *unchallenged* cases. As I have already pointed out, the exploration of "abnormal" cases is only intended as a means of dispelling certain misrepresentations concerning the normal cases. Deciding whether a concept can legitimately be applied to a new set of circumstances involves sorting out the similarities between the circumstances in question and other circumstances to which we have previously applied that concept. However, what is to count as "similar" or "different" is not a self-transparent idea. As Brian Bix puts it:

We are often willing to describe different objects or slightly different situations as 'the same'. How much leeway there is in the expression 'the same' (that is, how different objects or situations have to be before we feel compelled to describe them as 'different') can thus be seen to be itself responsive to, and reflective of, our usages and practices. (Bix, 1995: 148)

By contrast, Waismann's discussion of open texture focuses on how we apply our concepts in *unusual* circumstances. The novel circumstances envisaged by Waismann comprise totally new experiences, such as at present we cannot even imagine, and new discoveries "which would affect our whole interpretation of certain facts" (1945: 127). It is precisely in such unusual cases that our concepts would fail to apply unequivocally. Waismann does not seem to be interested in how the application of our concepts works in unchallenged cases, nor does he try to extrapolate his conclusions from one type of cases to the other.

Moreover, Waismann holds that open texture is rooted in the incompleteness of our empirical knowledge. In his view, it is because our knowledge can always be extended, and because some of our current beliefs might prove wrong, that any term's application might become indeterminate. This claim seems to be inconsistent with the outlook of the *Investigations*. As Bix argues (drawing on Gordon Baker's and Peter Hacker's famed commentary on the *Investigations*), one of Wittgenstein's main contentions is that in understanding language "we must distinguish the agreements that partly constitute the meaning of particular terms and the background conditions of the natural world [...]

which are 'the framework within which our language-games are played, not part of the games themselves'" (Bix, 1995: 16).<sup>5</sup> Bix goes on to stress that "Waismann's concept of 'open texture' arguably conflates the two, by claiming that changes in the background conditions might cause problems for the definition or application of terms" (1995: 16–7). To put the matter slightly differently, we would expect a proponent of Wittgenstein's views to insist that our uncertainty with regard to the application of our terms in unusual cases does not arise from the fact that in such cases, *as opposed to* familiar cases, rules do not uniquely determine the application of our terms. On this view, the difference between familiar cases and unfamiliar ones is to be accounted for by reference to our linguistic practice. By contrast, a proponent of Waismann's views would maintain that it is because our knowledge is limited that we cannot formulate *complete* rules which would provide for the application of our terms in unusual cases.

#### 3 Two readings of Waismann's notion of open texture

In what follows, I will distinguish between two readings of Waismann's notion of open texture. According to the weaker reading, holding that open texture is a pervasive feature of language comes down to holding that we cannot construct complete definitions of our terms, which would settle of an ultimate manner all possible questions concerning their future usage. On this reading, the real interest of Waismann's view lies in its emphasis on the fact that we cannot possibly anticipate all the circumstances in which our terms might be put into test, and consequently, we cannot come up in advance with all the necessary specifications. According to the stronger reading, to hold that our terms are open textured is to hold that even when we are confronted with *actual* cases in which our terms fail to apply unequivocally, we can expect such indeterminacy of application to persist. My aim here is not to reach a definitive conclusion about which of these readings is licensed by Waismann's writings. The distinction between the two above-mentioned readings will mainly serve as a background for discussing the issue of whether the picture of language advanced by metaphysical realists can accommodate the claim that most our terms are open textured.

Let me start by considering the weaker reading. According to it, the claim that open texture cannot be eliminated from our linguistic apparatus comes down to the claim that we can never eliminate the prospect that *new* cases will emerge in which our concepts, as previously defined, will fail to apply unequivocally. As Frederick Schauer emphasizes, potential vagueness cannot be eliminated "precisely because at a given point in time we cannot identify it" (1987: 31). However, "vagueness [...] once identified, is in theory eliminable, because by identifying the vagueness we have identified the way to provide the requisite specification" (ibid.).

This reading seems to be in line with Waismann's view according to which the process of defining and refining our concepts never reaches a final stage (1945: 125). It is also worth emphasizing that, according to him, another way of expressing the idea that open texture cannot be eliminated "would be to say that definitions of open terms are *always* corrigible or emendable" (1945: 123, emphasis in original). Waismann seems to

<sup>&</sup>lt;sup>5</sup> The quote comes from Baker and Hacker (1985: 229).

suggest that, once we encounter actual cases of indeterminacy, the means to remedy the indeterminacy will soon be at our disposal. In his *Principles of Linguistic Philosophy*, he points out that:

[W]e must admit that grammar is incomplete, and that should the circumstances arise we would make it more complete by introducing new rules to provide for such situations. (Waismann, 1965: 76)

A possible extension of Waismann's ideas about open texture is suggested by the following excerpt from Jerry Fodor<sup>6</sup>:

[T]o ask what we would say should certain of our current beliefs prove false involves asking what new beliefs we would then adopt. But to answer this question we would now have to be able to predict what theories would be devised and accepted were our current theories to prove untenable. Clearly, however, it is unreasonable to attempt to predict what theories would be accepted if our current theories were abandoned and, *a fortiori*, it is unreasonable to attempt to make such prediction on the basis of an appeal to our current linguistic intuitions. (Fodor, 1964: 207)

In other words, even if we can conceive of a number of cases in which objects behave abnormally, it is unreasonable to build decisions about such cases into the meaning of our terms. This interpretation shifts the accent from the idea that we cannot envisage beforehand all possible situations in which our terms might be put into test to the idea that our empirical terms cannot be precisely defined even with respect to all the situations that come to our mind. Arguably, it is the former idea that Waismann tried to convey by the metaphor of open texture. However, there is reason to believe that the latter idea is also consistent with his views on language. According to him, language is designed to respond to normal background conditions (1965: 76). As he emphasizes, "a concept is good if it fulfils the purpose for which it has been devised" (1965: 223). Waismann goes on to add that, for most of our practical aims, it suffices to delimit our concepts in several directions. What's more, it is not at all clear what purposes would be served by delimiting our concepts in many other possible directions and what are criteria that we would use in trying to accomplish such a task.

I turn now to the stronger reading of the notion of open texture. According to this reading, the claim that indeterminacy of the open texture variety cannot be eliminated should be taken to imply that, even when we encounter *actual* cases in which our terms fail to apply unequivocally, this indeterminacy of application will persist. There are two points one might wish to emphasize here. First, Waismann contends that even our most precise terms might, in unanticipated circumstances, become vague. To stick to one of his examples, suppose we came across a substance which had the chemical composition of gold, but emitted a new sort of radiation. Waismann points out that we would be puzzled about whether the term "gold" can legitimately be applied to this substance. Now, the fact that Waismann uses examples in which some of our most deeply entrenched beliefs seem to be violated suggests, *pace* Schauer, that even when we identify vagueness, we might have trouble coming up with the requisite specifications. In fact, one can argue that our

<sup>&</sup>lt;sup>6</sup> For this suggestion, see Margalit (1979).

puzzlement concerning the application of a natural kind term to an entity which seems to possess all the features that have hitherto been taken to be necessary and sufficient for kind membership, but which also possesses features which disqualify it from belonging to the kind in question, has to do with the fact that the reasons against applying the term seem to be just as strong as the reasons in favor of applying it. Still, the question remains whether claiming that we might have trouble formulating rules that would eliminate the indeterminacy is actually consistent with Waismann's view on language. Waismann, like Wittgenstein, rejects the picture of language advanced by metaphysical realists. Yet, as Bix points out, there are reasons to doubt that the above-mentioned claim is consistent with a picture of language that equates meaning with use. Bix draws attention to the fact that those who maintain that the use of a term exhausts its meaning will have a hard time explaining why changes in the background conditions of the natural world might cause difficulties for the definitions of our terms. However, my primary interest here is in whether the picture of language advanced by metaphysical realists can accommodate the claim that open texture is a pervasive feature of language. So, interesting as it might be, I will not pursue Bix's line of thought any further, since this move would involve an unnecessary detour from my main theme.

Second, one can argue that, if we focus on the view that the meaning of a term is dictated by its use, there is a clear sense in which indeterminacy can be said to persist. On this view, the application of our terms becomes indeterminate in unanticipated situations precisely because their previous use offers no clear guidance about the correct way to go on. A corollary of this view is that identifying indeterminacy does not automatically bring with it the means to remedy it. If the meaning of a term is exhausted by the way we use it, then it seems that indeterminacy of the open texture variety will persist until the new use of a term becomes firmly established. For example, in order to claim that a certain natural kind term applies to a previously unconsidered case, it does not suffice that an expert (i.e. a scientist) maintains that it does. What is needed is that the new use of the term becomes well-established among ordinary speakers. In short, even if we assume that modifying or extending the previous definitions of our term raises no special difficulties, it remains the case that indeterminacy cannot be eliminated right away.

At this point, one can emphasize that the two readings of the notion of open texture discussed here do not necessarily exclude one another. One might, for instance, read Waismann as holding both that (a) *potential* vagueness, as opposed to *actual* vagueness, cannot possibly be eliminated from our linguistic apparatus, and that (b) although when confronted with actual cases in which the application of our terms becomes vague we will eventually formulate rules that would remedy the indeterminacy, establishing the new use of a term within a certain linguistic community might actually require time. Nonetheless, keeping these two readings apart is important for the following reason. One worry about the view that open texture is a fundamental feature of our terms is that it is supported by a specific picture of language, according to which meaning is intimately connected with use. It is not entirely obvious, for instance, whether the picture of language advanced by metaphysical realists can accommodate the view that most of our terms are open textured. Yet, some might argue that how one answers the question of whether or not this is the case will depend on which reading of the notion of open texture one has in mind. Let me detail.

Admittedly, as long as we focus on the weaker reading of the notion of open texture there is no reason to suppose that metaphysical realists would reject the view that most of our terms are open textured. If what is meant by claiming that open texture is a pervasive feature of language is that we cannot construct complete definitions of our terms, which would settle all possible questions concerning their use, then metaphysical realists would most likely endorse this claim. According to them, the application of at least some of our terms (i.e. natural kind terms) is answerable to the outside world. The definitions of these terms are not supposed to be static; they are supposed to evolve as our knowledge of the world evolves. Holding otherwise would seem to go against the grain of the metaphysical realist's view of language. It is worth stressing that metaphysical realists would not only endorse the claim that the definitions of our terms are always corrigible, but would also endorse one other claim that I discussed above. Recall that our being unable to construct exhaustive definitions of our terms can be traced back to two factors. On the one hand, we cannot possibly foresee all the circumstances in which our terms might be put into test. But on the other hand, even if we can conceive of a number of such circumstances, it seems unreasonable to build decisions about all such cases into the meaning of our terms. Metaphysical realists would undoubtedly agree that our terms cannot be precisely defined even with respect to all the situations that come to our mind. As we have seen already, one of the reasons why it is unreasonable to settle in advance certain questions regarding the applicability of our terms is that doing so would require making decisions about what beliefs we would adopt if our current theories proved wrong, and making such decisions on the basis of our current linguistic intuitions does not seem justified. Yet surely, anyone who lays emphasis on the contribution of the world to the extension of our terms would insist that certain questions concerning the applicability of our terms cannot be settled solely on the basis of our linguistic intuitions.

Some might argue that, if the stronger reading of the notion of open texture is taken as relevant, then metaphysical realists would reject the view that most of our terms are open textured. This argument would go as follows. On a picture of language that equates meaning with use, the application of our terms becomes indeterminate in unusual cases precisely because their previous use offers no clear guidance about the correct way to go on. Yet, on the picture of language advanced by metaphysical realists, even if we were uncertain about whether a given term applies in a particular case, it would be misleading to say that the application of the term is indeterminate. According to the causal theory of reference, determining the meaning of a term is amenable to empirical investigation. The proponents of the causal theory stress that the correct definition of a term like "gold" or "whale" is an *a posteriori* matter (or, to put it differently, the question of how to define terms like "gold" or "whale" cannot be settled by means of conceptual analysis). Insofar as determining the meaning of such terms involves investigating the features of the world, and insofar as their definitions are essentially revisable – the argument would go – the fact that nothing in the previous use of a term dictates whether the term can properly be applied to a novel case is insufficient to establish that its application is indeterminate.

In the following section, I will attempt to prove that, despite common assumptions to contrary, the metaphysical realist's picture of language can accommodate the stronger reading of the claim that open texture is a fundamental characteristic of our terms. More specifically, I will offer a brief sketch of how a causal theory of reference combined with a property-cluster account of natural definitions like the one advocated by Richard Boyd

allows metaphysical realists to embrace the view that indeterminacy in the extension of some of our terms is real and ineliminable.

### 4 Homeostatic property-cluster definitions

The causal theory of reference was offered as an alternative to the traditional descriptivist theory of reference.<sup>7</sup> According to the latter, the reference of a term (e.g. a proper name) is fixed by a description (or a cluster of descriptions). Yet, as Saul Kripke (1972) rightly emphasized, a proper name can successfully refer to an individual even if the description associated with the name in question turns out to be false, or even if the speaker cannot come up with a description that is uniquely satisfied by that particular individual. Kripke held that the reference of a proper name is fixed by a causal-historical chain which traces back the use of a term to an initial act of "baptism". For instance, our current uses of the name "Aristotle" refer to the famous Greek philosopher Aristotle because they are part of a causal-historical chain that begins with those who originally gave Aristotle his name. This causal chain continues with various speakers "borrowing reference" from earlier speakers. On Kripke's view, provided that the right kind of causal relation is preserved, a speaker can use the name "Aristotle" to refer to Aristotle even if the description he associates with the name is false, or even if he cannot come up with a description that is uniquely true of the referent.

According to the proponents of the causal theory of reference, natural kind terms function partly like proper names, i.e. natural kind terms acquire their reference via an initial baptism, when a term is attached to an object (by ostension or by a description<sup>8</sup>), and then the term is passed on from one speaker to another in a reference-preserving causal-historical chain. Adopting a causal theory of reference for natural kind terms has the advantage of explaining how reference is possible prior to knowledge of the exact nature of the object being referred to. Undoubtedly, our initial beliefs about what exactly is required for membership in a natural kind may be erroneous. To take a commonly used example, the superficial properties by which we originally identified gold turned out to be irrelevant in determining whether or not a particular substance is gold. Yet, if natural kind terms are introduced ostensively, then reference is possible prior to the discovery of the properties that would need to figure in an adequate definition of a natural kind term. According to the proponents of the causal theory, even though our understanding of the nature of the object to which a term refers may drastically change, reference remains fixed. This continuity of reference is what allows us to conceive of competing scientific theories as genuinely conflicting. If the reference of terms like "water" or "electrical charge" is not determined by the descriptions associated with them, then rival scientists can be seen as disagreeing with one another instead of just talking past one another. The

<sup>&</sup>lt;sup>7</sup> The causal theory of reference was initially put forward by Kripke (1972) and Putnam (1973; 1975*a*; 1975*b*; 1975*c*). For a further elaboration, see Boyd (1993).

<sup>&</sup>lt;sup>8</sup> Kripke concedes that in some cases the initial baptism may involve a description. However, he contends that in such cases, contrary to what descriptivist theorists believe, the description is not giving a synonym of the name. He argues that the description merely "fixes the reference by some contingent marks of the object", and that "the name denoting that object is then used to refer to that object, even in referring to counterfactual situations where the object doesn't have the properties in question" (Kripke, 1972: 309).

causal theory of reference would thus vindicate our intuitive view of scientific progress. On the one hand, there seems to be no reason to worry that scientists are in fact referring to different realities despite using common terminology, so past scientific research can be seen as continuous with present research. On the other hand, if the initial use of a term is causally connected with the object being referred to, it is reasonable to assume that our knowledge of the nature of the object evolves as the causal give-and-take with the object goes on.<sup>9</sup> In other words, it is reasonable to assume that scientific theories which replace their predecessors provide an improved understanding of the phenomena they study.

As Boyd emphasizes, the causal theory of reference accords "with Quinean dicta to the effect that there are no analytic definitions, no truths by conventions" (1993: 491). Advocates of the causal theory point out that it was an empirical discovery that gold is the element with atomic number 79. It follows then that statements about what is required for membership in a natural kind are not analytic or *a priori*. It is important to stress that the traditional descriptivist theory of reference cannot accommodate the idea that what is required for membership in a natural kind is a matter for empirical investigation. This is because, on the aforementioned view, the reference of a natural kind term is fixed by means of a description which is analytically tied to the term. The awkward conclusion that we know everything there is to know about kind membership without investigating the features of the world is therefore unavoidable. In fact, as Christopher Hughes rightly points out, "there is a certain irony in the fact that traditionally empiricists have embraced the sort of pure descriptivist account of natural kind terms that makes knowledge of (scientific) kinds a non-empirical, a priori matter" (2004: 55, n80).<sup>10</sup>

The two above-mentioned theories of reference provide different approaches to linguistic precision. As Boyd argues:

The empiricist account readily supplies criteria of linguistic precision. Two uses of the same term (or two lexicographically different terms for that matter) are coreferential or co-extensive only when they are governed by the same definitional conventions. Vagueness arises from inexplicitness or intersubjective variation in definitional conventions, and ambiguity from the association of a single term with two or more nonequivalent definitional conventions. Both sources of linguistic imprecision have the same remedy: each general term should be associated with a single, quite explicit, and definite conventional definition which is accepted by the relevant linguistic community *prior to* the employment of the term in question. Linguistic precision can be identified with the existence of explicit, detailed, and intersubjectively accepted conventional definitions. (Boyd, 1993: 497, emphasis in original)

<sup>&</sup>lt;sup>9</sup> According to some proponents of the causal theory of reference, the referent must figure in the causal history of the introduction of the name. See Boyd (1993). Boyd claims that the causal theory of reference is correct precisely because the causal theory of knowledge is correct (1993: 503). Compare Kripke (1972). For pointing out some of the difficulties involved in holding that we can only introduce names for things that have had a causal impact on us, see Hughes (2004: 43–6).

<sup>&</sup>lt;sup>10</sup> See also Boyd (1993), for arguing that the empiricist account of natural kind terms goes against the grain of realism. As Boyd puts it, "the extensions of the 'natural' kinds [...] are – according to the empiricist conception – largely fixed by arbitrary and empirically unrevisable definitional conventions" (1993: 497–8).

Boyd goes on to emphasize that, according to the causal theory of reference, there is no such thing as a distinctly linguistic precision (1993: 500, 521-3). On Boyd's version of the causal theory, our linguistic categories must correspond to causally and explanatorily significant features of the world. Given that epistemic success is linguistically mediated, the accommodation of language to the causal structure of the world is crucial (1993: 483, 517). Boyd points out that it is rather doubtful that each of the general terms we currently employ is successful in "cutting nature at its joints" (1993: 521). Recall that, for instance, we used to classify whales under the term "fish" until it was discovered that whales are marine mammals. Or think about how the term "element" was used before the discovery of isotopes. Boyd stresses that scientific discoveries such as the ones mentioned above call for a modification in the use of the existing terminology and for the introduction of a new terminology in order to achieve the requisite accommodation between language and the relevant causal structures. Typically, there is more than one way to achieve such an accommodation (1993: 513-5; 1999: 77-9). For instance, when it was discovered that chemical elements can exist in more than one form, scientists could have introduced the term "isotope" (as they in fact did) to refer to atoms of the same chemical element which have a different number of neutrons, but they could have just as well reserved the term "element" for the categories we now call isotopes and introduced a different terminology for the resulting new elements. As Boyd argues, the important thing was that a change in our linguistic practice be effected.<sup>11</sup> It should be stressed, however, that, while the causal theory of reference allows for improvements in linguistic usage, such improvements have to do with the discovery of previously unknown causal features of the world. As Boyd puts it, precision in the use of scientific language is not "a matter of the proper following of linguistic rules", but rather "a matter of care in treating epistemological issues" (1993: 521).

Boyd further points out that, in contrast to the empiricist account, which places a special emphasis on linguistic precision, the causal account of reference requires that the definitions of natural kinds reflect a deference to the world on our part, even at the cost of higher conceptual complexity. He argues that, for instance, the remedy to the linguistic imprecision which manifests itself as vagueness is, in most cases, the epistemological precision which at the end of the day results in the resolution of natural kind terms is always indicative of some sort of imprecision. Boyd maintains that, quite to the contrary, indeterminacy in the application of some natural kind terms is indicative of precision in the accommodation of language to causally significant features of the world (1993: 530). Here is where the idea of homeostatic property-cluster kinds comes into the picture. Let me explain.

Boyd rejects the common view according to which the definitions of natural kinds must specify necessary and sufficient conditions for kind membership, and claims that there is an entire class of scientifically important kinds, which he refers to as *homeostatic property-cluster kinds*, for which the above-mentioned conception of natural definitions fails even as an idealization (1993: 484). He contends that homeostatic property-cluster

<sup>&</sup>lt;sup>11</sup> Boyd points out that the discovery of isotopes called for a refinement of the earlier use of the term "element". It is worth emphasizing that such a denotational refinement would be ruled out by the traditional descriptivist theory, given that, on this view, the reference of our terms is fixed once and for all by means of a description.
terms "have definitions which are provided by a collection of properties such that the possession of an adequate number of these properties is sufficient for falling within the extension of the term" (1988: 196). Yet, he insists that there will be some indeterminacy in the extension of such terms, and that this indeterminacy cannot be remedied by coming up with more precise definitions.

One important difference between the definitions of property-cluster kinds and the definitions envisaged by empiricists is that the properties that are part of the definition of a homeostatic property-cluster kind "are united *causally rather than conceptually*" (1993: 484, emphasis in original). The idea that a group of properties are causally united is spelled out in passage where Boyd argues that it is not uncommon the following sort of situation occurs:

1. There is a family F of properties that are 'contingently clustered' in nature in the sense that they co-occur in an important number of cases.

2. Their co-occurrence is not, at least typically, a statistical artifact, but rather the result of what may be metaphorically (sometimes literally) described as a sort of *homeostasis*. Either the presence of some of the properties in F tends (under appropriate conditions) to favor the presence of the others, or there are underlying mechanisms or processes which tend to maintain the presence of the properties in F, or both.

3. The homeostatic clustering of the properties in F is causally important: there are (theoretically or practically) important effects which are produced by a conjoint occurrence of (many of) the properties in F together with (some or all of) the underlying mechanisms in question. (Boyd, 1988: 197, emphasis in original)

Boyd further emphasizes that there are cases of imperfect homeostasis, i.e. cases in which a thing displays some but not all of the properties in F (in such cases, only some of the relevant homeostatic mechanisms are at work). He argues that the terms that refer to such homeostatic phenomena do not have analytic definitions. In other words, the question of which properties and mechanisms are part of the definition of a homeostatic propertycluster kind is not a question of conceptual elucidation, but a question that can only be answered a posteriori. Yet, given the possibility of imperfect homeostasis, it is likely that there will be some indeterminacy in extension associated with such homeostatic propertycluster definitions. As already pointed out, the relative importance of various properties and mechanisms in determining whether a property-cluster term can properly be applied to a certain thing is to be determined on *a posteriori* grounds. Boyd stresses that such a determination characteristically depends on how the underlying homeostatic mechanisms operate and on the causal efficiency of a given *partial* property-cluster. Still, there may be cases when a thing displays some but not all of the properties in F and when rational considerations do not univocally determine whether it falls within the extension of the property-cluster term in question. Boyd insists that it would be a mistake to believe that indeterminacy of this sort can be remedied by supplying more precise definitions. He argues that such definitional attempts would fail to preserve the naturalness of the kinds in question by requiring that "we treat as important distinctions which are irrelevant to causal explanation or to induction or that we ignore similarities which are important in just these ways" (1988: 198). In Boyd's view, precision in defining homeostatic propertycluster kinds does not consist in supplying definitions that would eliminate all extensional vagueness. *Epistemological* (as opposed to purely *linguistic*) precision requires that we pay heed to the definitional complexity of such kinds. More specifically, when it comes to borderline cases, precision requires simply that we take notice of the ways in which they are similar to and of those in which they differ from typical members of the kind (1988: 199).

In short, we have seen that a causal theory of reference combined with a propertycluster account of natural definitions allows metaphysical realists to endorse the claim that indeterminacy in the extension of some natural kind terms is real and ineliminable. According to Boyd, linguistic accommodation to causally and explanatorily significant features of the world requires that we abandon the idea that natural kinds must be defined in terms of necessary and sufficient conditions. Boyd claims that there is a whole class of scientifically important kinds whose natural definitions involve a property-cluster. As he emphasizes:

The property-cluster form of such definitions and the associated indeterminacy are dictated by the scientific task of employing categories which correspond to inductively and explanatorily relevant causal structures. In particular, the indeterminacy in extension of such natural definitions could not be remedied without rendering the definitions *un*natural in the sense of being scientifically misleading. (Boyd, 1998: 196, emphasis in original)

The upshot of all these considerations is that metaphysical realists have no trouble accommodating the stronger reading of the claim that open texture is a pervasive feature of language. This is important because it helps to dispel a worry that was raised in section 2.6.3. The worry can be expressed as follows. Some might argue that the thesis according to which in unanticipated circumstances the application of our terms might become indeterminate is licensed by a particular view of language, one that metaphysical realists rightfully reject. This, in turn, would cast doubt on whether moral realists can consistently endorse the claim that the difficulties we sometimes experience in determining whether a moral term applies to a novel case are indicative of the fact that there is no determinate answer to whether the term in question can legitimately be applied. In this appendix, I have pointed out that a causal or naturalistic theory of reference is consistent with the view that indeterminacy in the application of our natural kind terms is real and ineliminable. As Boyd argues, such indeterminacy is a necessary result of "cutting nature at its joints".

Before concluding, let me forestall a possible objection. The argument developed here shows that a causal or naturalistic theory of reference is consistent with the view that indeterminacy in the application of some of our terms is real and ineliminable. This in turn is supposed to allay the suspicion that moral realists would have trouble accommodating indeterminacy. Yet, one might wonder whether a causal moral semantics is plausible and whether this affects the present argument in any way. Boyd, for instance, maintains that the causal theory of reference applies not only to natural kind terms, but also to moral terms. On this view, the reference of a moral term is established by causal connections of the right sort between the use of a term and its referent. However, as Terry Horgan and Mark Timmons have recently pointed out, there are doubts about whether a causal moral semantics is defensible.<sup>12</sup> Horgan and Timmons ask us to imagine a "Moral

<sup>&</sup>lt;sup>12</sup> See Horgan and Timmons (1991; 1992; 2000).

Twin Earth" whose inhabitants use the term "right" in much the same way we do. Recall that, according to Boyd, a moral term refers to the property that causally regulates our use of that term. Horgan and Timmons ask us to suppose that investigation into Twin Earth moral discourse and practice would reveal that the use of the term "right" on Twin Earth is causally regulated by a natural property that is distinct from the one that regulates the use of the term on Earth. Thus, the proponent of Boyd's view would have to conclude that any moral disagreement between Earthlings and Twin Earthlings is merely *apparent* (since moral terms and twin moral terms refer to different properties). And surely, this conclusion is counterintuitive, since the inhabitants of Moral Twin Earth use moral terms the same way as we do.

In response, let me first emphasize that my aim here has been a modest one, i.e. to show that one cannot object to the view that the application of our moral terms might be become indeterminate on the grounds such a view is committed to the wrong semantic theory. (Waismann, who was the first to argue that the application of *any* term, no matter how precise, can be become indeterminate in unanticipated circumstances, was a Wittgensteinian scholar. So the question naturally arises whether the claim that open texture is a pervasive feature of language makes sense only on a certain semantic theory.) I have argued that the view according to which the application of our terms can become indeterminate is compatible with more than one semantic theory. More concretely, I have shown that the causal theory of reference, which is typically adopted by metaphysical realists, can accommodate this view. My argument implies nothing about whether moral realists would have to adopt a causal semantics.

Still, a careful reader might point out that, in reaching the conclusion that the causal theory of reference can allow for indeterminacy, I have relied on Boyd's version of the causal theory, and that the idea that realism cannot be divorced from a causal semantics is part and parcel of Boyd's theory. Thus, one might argue that I have committed myself to the view that moral realists have to embrace a causal semantics.

Boyd holds that the empiricist account of general terms goes against the grain of realism. As he puts it, "the extensions of the 'natural' kinds, or values of the magnitudes which are the referents of general terms are – according to the empiricist conception – largely fixed by arbitrary and empirically unrevisable definitional conventions" (1993: 497–8). He claims that a causal, nondefinitional, account of reference fixing is correct for a wide range of general terms, including moral terms. Assessing the merits of Boyd's view about the semantics of moral terms would take me too far afield. It is important to stress, however, that even if turns out that a causal moral semantics is indefensible, this does not mean that we would be forced to adopt some form of moral antirealism (as both Boyd and his critics would have us believe). While acknowledging that the Moral Twin Earth scenario proposed by Horgan and Timmons poses a serious challenge to the sort of moral semantics that Boyd defends, several authors have attempted to prove that the idea that moral terms admit of synthetic definitions can be salvaged.<sup>13</sup> If it is possible to steer

<sup>&</sup>lt;sup>13</sup> For arguments to the effect that the Moral Twin Earth objection does not undermine the view that moral terms have synthetic definitions, see Sayre-McCord (1997), Brink (2001), and Van Roojen (2006). Such arguments aim to prove that a modified version of Boyd's moral semantics, which downplays the role of causal regulation, can handle the aforementioned objection. Van Roojen, for instance, claims that the kind of regulation that is involved in Boyd-style moral semantics should be modified from *causal* to *epistemic*, while Brink argues that the objection raised by Horgan and Timmons loses its bite once we understand

clear of the view that the reference of our moral terms is determined once and for all by arbitrary definitional conventions, by showing for instance that our use of such terms is responsive to advancements in moral knowledge, then there seems to be no reason to complain that moral terms do not genuinely refer. However, a detailed discussion of these issues lies beyond the scope of this dissertation.

reference in terms of referential intentions rather than in terms of causal regulation. For other replies to the Moral Twin Earth objection, see Copp (2000) and Merli (2002).

## **Bibliography**

- Ackerman, Felicia. 1994. Roots and Consequences of Vagueness. *Philosophical Perspectives* 8: 129–136.
- Alston, William P. 1964. Philosophy of Language. Englewood Cliffs, N.J.: Prentice-Hall.
- 1967. Vagueness. In Paul Edwards (ed.), *The Encyclopedia of Philosophy*, Vol. 8. New York: Macmillan, 218–221.

Audi, Robert. 1997. Ethical Naturalism and the Explanatory Power of Ethical Concepts. In Robert Audi, *Moral Knowledge and Ethical Character*. Oxford: Oxford University Press, 112–128.

— 2002. Prospects for a Value-Based Intuitionism. In Philip Stratton-Lake (ed.), *Ethical Intuitionism: Re-evaluations*. Oxford: Oxford University Press, 29–55.

Ayer, A. J. 1936. Language, Truth and Logic. London: Gollancz.

- Baker, G. P. and Hacker, P. M. S. 1984. Scepticism, Rules and Language. Oxford: Blackwell.
- 1985. Wittgenstein: Rules, Grammar and Necessity. An Analytical Commentary on the "Philosophical Investigations", Volume 2. Oxford: Blackwell.
- Berlin, Isaiah. 1959. The Pursuit of the Ideal. In Isaiah Berlin, *The Crooked Timber of Humanity: Chapters in the History of Ideas*. Princeton, New Jersey: Princeton University Press, 1–19.

Bix, Brian. 1995. Law, Language and Legal Determinacy. Oxford: Clarendon Press.

Blackburn, Simon. 1971. Moral Realism. In John Casey (ed.), *Morality and Moral Reasoning*. London: Methuen, 101–124.

— 1985*a*. Errors and the Phenomenology of Value. In Ted Honderich (ed.), *Morality and Objectivity*. London: Routledge & Kegan Paul, 1–22.

— 1985*b*. Supervenience Revisited. In Ian Hacking (ed.), *Exercises in Analysis*. Cambridge: Cambridge University Press, 47–67.

- 1992. Through Thick and Thin. *Proceedings of the Aristotelian Society*, Suppl. Vol. 66: 285–299.
- 1993. Essays in Quasi-Realism. New York: Oxford University Press.
- Boot, Martijn. 2009. Parity, Incomparability and Rationally Justified Choice. *Philosophical Studies* 146: 75–92.

Boyd, Richard. 1988. How to Be a Moral Realist. In Geoffrey Sayre-McCord (ed.), *Essays on Moral Realism.* Ithaca, NY: Cornell University Press, 181–228.

- 1993. Metaphor and Theory Change. In Andrew Ortony (ed.), *Metaphor and Thought*. New York: Cambridge University Press, 481–532.
- 1999. Kinds as the "Workmanship of Men": Realism, Constructivism, and Natural Kinds. In Julian Nida-Rümelin (ed.), *Rationalität, Realismus, Revision: Proceedings of the Third International Congress, Gesellschaft für Analytische Philosophie*. Berlin: de Gruyter, 52–89.
- 2006. Homeostasis, Species, and Higher Taxa. In Marc Lange (ed.), *Philosophy of Science: An Anthology*. Oxford: Blackwell, 406–426.
- Brennan, J. M. 1977. *The Open Texture of Moral Concepts*. New York: Barnes and Noble Books.

Brink, David O. 1984. Moral Realism and the Sceptical Arguments from Disagreement and Queerness. *Australian Journal of Philosophy* 62: 111–125.

— 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

— 1996. Moral Conflict and Its Structure. In Homer. E. Mason (ed.), Moral Dilemmas and Moral Theory. New York & Oxford: Oxford University Press, 102–126.

— 2001. Realism, Naturalism, and Moral Semantics. In Ellen Frankel Paul, Fred D. Miller, and Jeffrey Paul (eds.), *Moral Knowledge*. Cambridge: Cambridge University Press, 154–176.

Broome, John. 1992. Review of Isaac Levi's *Hard Choices*. *Economics and Philosophy* 8: 169–176.

— 1997. Is Incommensurability Vagueness? In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 67–89.

— 1999. Incommensurable Values. In John Broome, *Ethics Out of Economics*. Cambridge: Cambridge University Press, 145–161.

Carlson, Erik. 2004. Broome's Argument Against Value Incommensurability. *Utilitas* 16: 92–98.

Cartwright, Nancy. 1983. How the Laws of Physics Lie. Oxford: Oxford University Press.

Chang, Ruth. 1997. Introduction. In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 1–34.

- 2001. Against Constitutive Incommensurability or Buying and Selling Friends. *Philosophical Issues* 11: 33–60.
- 2002a. Making Comparisons Count. New York: Routledge.
- 2002b. The Possibility of Parity. Ethics 112: 659–688.
- 2004. Putting Together Morality and Well-Being. In Peter Baumann and Monika Betzler (eds.), *Practical Conflicts: New Philosophical Essays*. Cambridge: Cambridge University Press, 118–158.
- Conee, Earl. 1982. Against Moral Dilemmas. The Philosophical Review 91: 87-97.
- 1989. Why Moral Dilemmas Are Impossible. *American Philosophical Quarterly* 26: 133–141.
- Copp, David. 1991. Moral Realism: Facts and Norms. Ethics 101: 610-624.
  - 2000. Milk, Honey, and the Good Life on Moral Twin Earth. *Synthese* 124: 113–137.
- Cuneo, Terence. 2007*a. The Normative Web An argument for Moral Realism.* Oxford: Oxford University Press.

— 2007b. Moral Facts as Configuring Causes. In Terence Cuneo and Russ Shafer-Landau (eds.), *Foundations of Ethics: An Anthology*. Oxford: Blackwell, 353–365.

Dancy, Jonathan. 1986. Two Conceptions of Moral Realism. *Proceedings of the Aristotelian Society*, Suppl. Vol. 60: 167–187.

- 1993. Moral Reasons. Oxford: Blackwell Publishers.
- —— 1996. In Defense of Thick Concepts. *Midwest Studies in Philosophy* 20: 263–279.
- 2004. Ethics Without Principles. Oxford: Clarendon Press.

Davidson, Donald. 1986. On the Very Idea of a Conceptual Scheme. In Donald Davidson, *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press, 5–20.

Donagan, Alan. 1984. Consistency in Rationalist Moral Systems. *The Journal of Philosophy* 81: 291–309.

- 1996. Moral Dilemmas, Genuine and Spurious: A Comparative Anatomy. In Homer. E. Mason (ed.), *Moral Dilemmas and Moral Theory*. New York, Oxford: Oxford University Press, 11–22.
- Dreier, James. 1992. The Supervenience Argument against Moral Realism. Southern Journal of Philosophy 30: 13–38.
- Dummett, Michael. 1978a. Preface. In Michal Dummett, *Truth and Other Enigmas*. London: Duckworth, ix-li.
- 1978b. Realism. In Michal Dummett, *Truth and Other Enigmas*. London: Duckworth, 145–165.
- 1981. Realism. In *The Interpretation of Frege's Philosophy*. London: Duckworth.
- 1982. Realism. *Synthese* 52: 55–112.
- 1993. Realism and Anti-Realism. In Michael Dummett, *The Seas of Language*. Oxford: Clarendon Press, 462–78.
- Dupré John. 1981. Natural Kinds and Biological Taxa. Philosophical Review 90: 66-90.
- Dworkin, Ronald. 1996a. Indeterminacy and Law. In Stephen Guest (ed.), *Positivism Today*. Brookfield, VT: Dartmouth, 1–9.
- 1996b. Objectivity and Truth: You'd Better Believe It. *Philosophy and Public Affairs* 25: 87–139.
- 1996c. Do Liberty and Equality Conflict? In Paul Barker (ed.), *Living as Equals*. Oxford: Oxford University Press, 39–57.
- 2001. Do Liberal Values Conflict? In Mark Lilla, Ronald Dworkin and Robert Silvers (eds.), *The Legacy of Isaiah Berlin*. New York: New York Review of Books, 73–90.
- Erikson, Erik. 1963. Childhood and Society. New York: Norton.
- Finnis, John. 1980. Natural Law and Natural Rights. Oxford: Clarendon Press.
- Firth, Roderick. 1952. Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research* 12: 317–345.
- Fodor, Jerry A. 1964. On Knowing What We Would Say. *The Philosophical Review* 73: 198–212.
- Foot, Philippa. 1983. Moral Realism and Moral Dilemma. *The Journal of Philosophy* 80: 379–398.
  - 2002. Moral Dilemmas Revisited. In Philippa Foot, *Moral Dilemmas and Other Topics in Moral Philosophy*. Oxford: Clarendon Press, 175–188.
- Gert, Bernard. 2005. *Morality: Its Nature and Justification*. Oxford: Oxford University Press.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment.* Cambridge, Mass.: Harvard University Press.
- Gowans, Christopher. 1994. Innocence Lost: An Examination of Inescapable Moral Wrongdoing. New York, Oxford: Oxford University Press.
- 2000. Introduction. In Christopher Gowans (ed.), *Moral Disagreements: Classic and Contemporary Readings*. London & New York: Routledge, 1–43.

Griffin, James. 1977. Are There Incommensurable Values? *Philosophy and Public Affairs* 7: 39–59.

— 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press.

— 1991. Mixing Values. Proceedings of the Aristotelian Society, Suppl. Vol. 65: 101– 118.

— 1997. Incommensurability: What's the Problem? In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 35–51.

— 1998. *Value Judgement: Improving Our Ethical Beliefs*. Oxford: Oxford University Press.

— 2000. Replies. In Roger Crisp and Brad Hooker (eds.), *Well-Being and Morality: Essays in Honour of James Griffin*. Oxford: Oxford University Press, 281–313.

Guttenplan, Samuel. 1979–80. Moral Realism and Moral Dilemmas. *Proceedings of the Aristotelian Society* 80: 61–80.

Hale, Bob. 1997. Realism and Its Oppositions. In Bob Hale and Crispin Wright (eds.), *A Companion to the Philosophy of Language*. Oxford: Blackwell, 271–308.

- Hare, R. M. 1963. Freedom and Reason. Oxford: Clarendon Press.
- Harman, Gilbert. 1965. The Inference to the Best Explanation. *The Philosophical Review* 74: 88–95.
- 1973. *Thought*. Princeton, New Jersey: Princeton University Press.
- 1975. Moral Relativism Defended. *The Philosophical Review* 84: 3–22.
- —— 1977. The Nature of Morality. New York: Oxford University Press.
- —— 1986. Moral Explanations of Natural Facts Can Moral Claims Be Tested Against Moral Reality? *Southern Journal of Philosophy* 24 (Suppl.): 57–68.
- —— 1996. Moral Relativism. In Gilbert Harman and Judith Jarvis Thomson, *Moral Relativism and Moral Objectivity*. Oxford: Blackwell, 3–64, 157–187.

Hart, H. L. A. 1961. The Concept of Law. Oxford: Clarendon Press.

Hookway, Christopher. 1986. Two Conceptions of Moral Realism. *Proceedings of the Aristotelian Society*, Suppl. Vol. 60: 189–205.

Horgan, Terry and Timmons, Mark. 1991. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research* 16: 447–465.

- 1992. Troubles on Moral Twin Earth: Moral Queerness Revived. Synthese 92: 221–260.
- 2000. Coping Out on Moral Twin Earth. Synthese 124: 139–152.
- 2006. Morality Without Moral Facts. In James Dreier (ed.), Contemporary Debates in Moral Theory. Oxford: Blackwell, 220–238.
- Hughes, Christopher. 2004. Kripke: Names, Necessity and Identity. Oxford: Oxford University Press.

Hurley, Susan. 1985. Objectivity and Disagreement. In Ted Honderich (ed.), *Morality* and Objectivity. London: Routledge & Kegan Paul, 54–97.

— 1989. *Natural Reasons*. Oxford: Oxford University Press.

Jackson, Frank. 2008. The Argument from the Persistence of Moral Disagreement. In Russ Shafer-Landau (ed.), Oxford Studies in Metaethics, Volume 3. Oxford: Oxford University Press, 75–86.

- Johnston, Mark. 1989. Dispositional Theories of Value. Proceedings of the Aristotelian Society, Suppl. Vol. 63: 139–174.
- Joyce, Richard. 2007. Joyce, Richard, "Moral Anti-Realism", *The Stanford Encyclopedia* of *Philosophy (Summer 2009 Edition)*, Edward N. Zalta (ed.), URL =

<a href="http://plato.stanford.edu/archives/sum2009/entries/moral-anti-realism/">http://plato.stanford.edu/archives/sum2009/entries/moral-anti-realism/</a>>.

Kekes, John. 1993. *The Morality of Pluralism*. Princeton & New Jersey: Princeton University Press.

Kis, János. 2002. Natural Duties Regarding Institutions. Manuscript.

— 2008. Politics as a Moral Problem. Budapest: Central University Press.

- Koslicki, Kathrin. 2008. Natural Kinds and Natural Kind Terms. *Philosophy Compass* 3: 789–802.
- Kripke, Saul. 1972. Naming and Necessity. In Donald Davidson and Gilbert Harman (eds.), *Semantics of Natural Language*. Reidel: Dordrecht, 253–355.
- —— 1982. Wittgenstein on Rules and Private Language: An Elementary Exposition. Cambridge, Mass.: Harvard University Press.
- Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions* (2<sup>nd</sup> edition). Chicago: Chicago University Press.
- LaPorte, Joseph. 2004. *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.
- Leiter, Brian. 2001. Moral Facts and Best Explanations. In Ellen Frankel Paul, Fred D. Miller, and Jeffrey Paul (eds.), *Moral Knowledge*. Cambridge: Cambridge University Press, 79–101.
- Levi, Isaac. 1986. *Hard Choices: Decision Making Under Unresolved Conflict.* Cambridge: Cambridge University Press.
- Lipton, Peter. 2004. Inference to the Best Explanation (2<sup>nd</sup> edition). London: Routledge.
- Loeb, Don. 1998. Moral Realism and the Argument from Disagreement. *Philosophical Studies* 90: 281–303.
- Lukes, Steven. 1990. Incommensurability in Science and Ethics. In Steven Lukes, *Moral Conflict and Politics*. Oxford: Clarendon Press, 33–49.
  - 1997. Comparing the Incomparable: Trade-offs and Sacrifices. In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 184–195.
- Lyons, David. 1976. Ethical Relativism and the Problem of Incoherence. *Ethics* 86: 107–121.
- Mackie, J. L. 1977. Ethics: Inventing Right and Wrong. New York: Penguin Books.

Macklem, Timothy. 2001. Choice and Value. Legal Theory 7: 1–34.

Marcus, Ruth Barcan. 1980. Moral Dilemmas and Consistency. The Journal of Philosophy 77: 121–136.

— 1996. More about Moral Dilemmas. In Homer. E. Mason (ed.), Moral Dilemmas and Moral Theory. New York & Oxford: Oxford University Press, 23–35.

- Margalit, Avishai. 1979. Open Texture. In Avishai Margalit (ed.), *Meaning and Use*. Reidel: Dordrecht, 141–151.
- Mason, Andrew. 1993. *Explaining Political Disagreement*. Cambridge: Cambridge University Press.
- McConnell, Terrance. 1993. Dilemmas and Incommensurateness. *The Journal of Value Inquiry* 27: 247–252.

— 1996. Moral Residue and Dilemmas. In Homer. E. Mason (ed.), Moral Dilemmas and Moral Theory. New York, Oxford: Oxford University Press, 36–47.

- McDowell, John. 1981. Non-Cognitivism and Rule-Following. In Steven Holzman and C. Leich (eds.), *Wittgenstein: To Follow a Rule*. London: Routledge & Kegan Paul, 141–162.
- 1985. Values and Secondary Qualities. In Ted Honderich (ed.), *Morality and Objectivity*. London: Routledge & Kegan Paul, 110–129.
- 1998*a*. Virtue and Reason. In John McDowell, *Mind, Value and Reality*. Cambridge, Mass.: Harvard University Press, 50–73.

— 1998b. Wittgenstein on Following a Rule. In John McDowell, *Mind, Value and Reality*. Cambridge, Mass.: Harvard University Press, 221–262.

McGinn, Colin. 1997. Ethics, Evil and Fiction. Oxford: Oxford University Press.

— 1999. Two Notions of Realism? In Colin McGinn, *Knowledge and Reality: Selected Essays*. Oxford: Oxford University Press, 267–78.

McGinn, Marie. 1997. Wittgenstein and the "Philosophical Investigations". London: Routledge.

Merli, David. 2002. Return to Moral Twin Earth. Canadian Journal of Philosophy 32: 207-240.

Miller, R. W. 1985. Ways of Moral Learning. The Philosophical Review 94: 507-556.

- Milo, Ronald. 1996. Contractarian Constructivism. *The Journal of Philosophy* 92: 181–204.
- Nagel, Thomas. 1979. The Fragmentation of Value. In Thomas Nagel, *Mortal Questions*. Cambridge: Cambridge University Press, 128–141.

— 1980. The Limits of Objectivity. In Sterling McMurrin (ed.), The Tanner Lectures

- on Human Values, Vol. 1. Salt Lake City: University of Utah Press, 75–139.
- 1986. The View from Nowhere. Oxford: Oxford University Press.

— 2001. Pluralism and Coherence. In Mark Lilla, Ronald Dworkin and Robert Silvers (eds.), *The Legacy of Isaiah Berlin*. New York: New York Review of Books, 105–111.

- Payne, Andrew. 2005. A New Account of Thick Concepts. *The Journal of Value Inquiry* 39: 89–103.
- Pettit, Philip. 1990. The Reality of Rule-Following. Mind 99: 1-21.
- 1991. Realism and Response-Dependence. *Mind* 100: 587–626.
- 2001. Embracing Objectivity in Ethics. In Brian Leiter (ed.), *Objectivity in Law and Morals*. Cambridge: Cambridge University Press, 234–286.

Plato. Euthyphro. 1997. In John M. Cooper (ed.), *Plato Complete Works*. Indianapolis: Hackett Publishing Company.

Putnam, Hilary. 1973. Meaning and Reference. The Journal of Philosophy 70: 699–711.

— 1975*a*. Is Semantics Possible? In Hilary Putnam, *Mind, Language and Reality*. *Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press, 139–152.

- 1975b. Explanation and Reference. In Hilary Putnam, *Mind, Language and Reality. Philosophical Papers, Volume 2.* Cambridge: Cambridge University Press, 196–214.
- 1975c. The Meaning of 'Meaning'. In Hilary Putnam, *Mind, Language and Reality. Philosophical Papers, Volume 2.* Cambridge: Cambridge University Press, 215–271.
- 2003. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, Mass.: Harvard University Press.

- Qizilbash, Mozaffar. 2000. Comparability of Values, Rough Equality, and Vagueness: Griffin and Broome on Incommensurability. *Utilitas* 12: 223–240.
- Quine, Willard Van Orman. 1969. Natural Kinds. In Willard Van Orman Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press, 114– 138.

Railton, Peter. 1986. Moral Realism. The Philosophical Review 95: 163-207.

— 1996. The Diversity of Moral Dilemma. In Homer. E. Mason (ed.), *Moral Dilemmas and Moral Theory*. New York, Oxford: Oxford University Press, 140–166.

—— 1998. Moral Explanation and Moral Objectivity. *Philosophy and Phenomenological Research* 58: 175–182.

Rawls, John. 1971. A Theory of Justice. Cambridge, Mass.: Harvard University Press.

— 1980. Kantian Constructivism in Moral Theory. *The Journal of Philosophy* 77: 515–572.

Raz, Joseph. 1985–86. Value Incommensurability: Some Preliminaries. *Proceedings of the Aristotelian Society* 86: 117–134.

- —— 1986. The Morality of Freedom. Oxford: Clarendon Press.
- —— 1989. Facing Up: A Reply. Southern California Law Review 62: 1153–1235.
- 1991. Mixing Values. Proceedings of the Aristotelian Society, Suppl. Vol. 65: 83– 100.

— 1997. Incommensurability and Agency. In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 110–128.

— 1999. The Truth in Particularism. In Joseph Raz, *Engaging Reason: On the Theory* of Value and Action. Oxford: Oxford University Press, 218–246.

— 2003. *The Practice of Value*. Oxford: Oxford University Press.

Regan, Donald. 1989. Authority and Value: Reflections on Raz's Morality of Freedom. Southern California Law Review 62: 995–1095.

— 1997. Value, Comparability and Choice. In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 129–150.

Richardson, Henry S. 1994. *Practical Reasoning about Final Ends*, Cambridge: Cambridge University Press.

Rosen, Gideon. 1995. The Shoals of Language. Mind 104: 599-609.

Ross, W. D. 1930. The Right and the Good. Oxford: Oxford University Press.

Sartre, Jean Paul. 1956. Existentialism is a Humanism. Trans. Philip Mairet. In Walter Kaufmann (ed.), *Existentialism from Dostoevsky to Sartre*. New York: Meridian Books, 287–311.

Sayre-McCord, Geoffrey. 1988a. The Many Moral Realisms. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Ithaca, NY: Cornell University Press, 1–23.

— 1988b. Moral Theory and Explanatory Impotence. In G. Sayre-McCord (ed.), *Essays on Moral Realism.* Ithaca, NY: Cornell University Press, 256–281.

- 1997. 'Good' on Twin Earth. *Philosophical Issues* 8: 267–292.
- 2005. Moral Realism. *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, Edward N. Zalta (ed.), URL =

<http://plato.stanford.edu/archives/win2009/entries/moral-realism/>.

- Scanlon, Thomas. 1982. Contractualism and Utilitarianism. In Amartya Sen and Bernard Williams (eds.), Utilitarianism and Beyond, Cambridge: Cambridge University Press, 103–128.
- Schaber, Peter. 2004. Are There Insolvable Moral Conflicts? In Peter Baumann and Monika Betzler (eds.), *Practical Conflicts: New Philosophical Essays*. Cambridge: Cambridge University Press, 279–294.
- Schauer, Frederick. 1987. Authority and Indeterminacy, *Nomos XXIX: Authority Revisited*. New York & London: New York University Press, 28–38.
- 1998. Instrumental Incommensurability. University of Pennsylvania Law Review 146: 1215–1233.
- Seung, T. K., and Bonevac, Daniel. 1992. Plural Values and Indeterminate Rankings. *Ethics* 102: 799–813.

Shafer-Landau, Russ. 1994. Ethical Disagreement, Ethical Objectivism and Moral Indeterminacy. *Philosophy and Phenomenological Research* 54: 331–344.

- 1995. Vagueness, Borderline Cases and Moral Realism. *American Philosophical Quarterly* 32: 83–96.
- 2003. Moral Realism: A Defence. Oxford: Oxford University Press.
- 2006. Ethics as Philosophy: A Defense of Ethical Nonnaturalism. In Terry Horgan and Mark Timmons (eds.), *Metaethics After Moore*. Oxford: Oxford University Press, 209–232.

Sinnott-Armstrong, Walter. 1988. Moral Dilemmas. Oxford: Basil Blackwell.

- 1996. Moral Dilemmas and Rights. In Homer. E. Mason (ed.), Moral Dilemmas and Moral Theory. New York, Oxford: Oxford University Press, 48–65.
- Skorupski, John. 1999. Ethical Explorations. Oxford: Oxford University Press.
  - 2000. Irrealist Cognitivism. In Jonathan Dancy (ed.), *Normativity*. Oxford: Blackwell, 116–139.

Smith, Michael. 1993. Objectivity and Moral Realism: On the Significance of the Phenomenology of Moral Experience. In John Haldane and Crispin Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press, 235–255.

- 1994. *The Moral Problem*. Oxford: Blackwell.
- 2000. Does the Evaluative Supervene on the Natural? In Roger Crisp and Brad Hooker (eds.), *Well-Being and Morality: Essays in Honour of James Griffin*. Oxford: Oxford University Press, 91–114.
- Sorensen, Roy A. 1988. Blindspots. Oxford: Clarendon Press.

— 1991. Moral Dilemmas, Thought Experiments, and Conflict Vagueness. *Philosophical Studies* 63: 291–308.

- Sosa, Ernest. 2001. Objectivity without Absolutes. In Alex Byrne, Robert Stalnaker and Ralph Wedgwood (eds.), *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson*. Cambridge, Mass.: MIT Press, 215–228.
- De Sousa, Ronald. 1974. The Good and the True. Mind 83: 534-551.
- Stevenson, Charles. 1937. The Emotive Meaning of Ethical Terms. Mind 46: 14-31.
- —— 1948. The Nature of Ethical Disagreement. *Sigma* 1–2: 8–9.
- Stocker, Michael. 1990. Plural and Conflicting Values. Oxford: Oxford University Press.
- Strandberg, Caj. 2008. Particularism and Supervenience. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Volume 3. Oxford: Oxford University Press, 129–158.

Sturgeon, Nicholas L. 1985. Moral Explanations. In David Copp and David Zimmerman (eds.), *Morality, Reason and Truth*. Totowa, N.J.: Rowman & Allanheld, 49–78.

— 1986*a*. Harman on Moral Explanations of Natural Facts. *Southern Journal of Philosophy* 24 (Suppl.): 69–78.

— 1986b. What Difference Does It Make Whether Moral Realism Is True? *Southern Journal of Philosophy* 24 (Suppl.): 115–141.

— 1994. Moral Disagreement and Moral Relativism. *Social Philosophy and Policy* 11: 80–115.

— 2002. Ethical Intuitionism and Ethical Naturalism. In Philip Stratton-Lake (ed.), *Ethical Intuitionism: Re-evaluations*. Oxford: Oxford University Press, 184–211.

— 2006. Moral Explanations Defended. In James Dreier (ed.), Contemporary Debates in Moral Theory. Oxford: Blackwell, 241–262.

Styron, William. 1980. Sophie's Choice. New York: Bantam Books.

Svavarsdóttir, Sigrún. 2001. Objective Values: Does Metaethics Rest on a Mistake? In Brian Leiter (ed.), *Objectivity in Law and Morals*. Cambridge: Cambridge University Press, 144–193.

Tersman, Folke. 2006. Moral Disagreement. Cambridge: Cambridge University Press.

Thomas, Alan. 2006. *Value and Context: The Nature of Moral and Political Knowledge*. Oxford: Oxford University Press.

Thomasson, Amie Lynn. 2003. Realism and Human Kinds. *Philosophy and Phenomenological Research* 67: 580–609.

— 2007. Ordinary Objects. Oxford: Oxford University Press.

Thomson, Judith Jarvis. 1996. Moral Objectivity. In Gilbert Harman and Judith Jarvis Thomson, *Moral Relativism and Moral Objectivity*. Oxford: Blackwell, 67–154, 188– 217.

Timmons, Mark. 1999. Morality Without Foundations: A Defense of Ethical Contextualism. New York, Oxford: Oxford University Press.

Tolhurst, William. 1987. The Argument from Moral Disagreement. *Ethics* 97: 610–621.

Unger, Peter. 1979. There Are No Ordinary Things. Synthese 41: 117-54.

- Van Roojen, Mark. 2006. Knowing Enough to Disagree: A New Response to the Moral Twin Earth Argument. In Russ Shafer-Landau (ed.), Oxford Studies in Metaethics, Volume 1. Oxford: Oxford University Press, 161–193.
- Väyrynen, Pekka. 2006. Moral Realism. In Donald Borchert (ed.), *Encyclopedia of Philosophy*, 2<sup>nd</sup> edition, Vol. 6. New York: Macmillan Reference USA, 379–382.

Waismann, Friedrich. 1945. Verifiability. *Proceedings of the Aristotelian Society*. Suppl. Vol. 19: 119–150.

— 1965. *The Principles of Linguistic Philosophy* (ed. Rom Harré). London: Macmillan Press.

Wasserman, Ryan. 2004. Indeterminacy, Ignorance and the Possibility of Parity. *Philosophical Perspectives* 18: 391–403.

Wellman, Carl. 1975. Ethical Disagreement and Objective Truth. American Philosophical Quarterly 12: 211–221.

White, Nicholas. 2004. Conflicting Values and Conflicting Virtues. In Peter Baumann and Monika Betzler (eds.), *Practical Conflicts: New Philosophical Essays*. Cambridge: Cambridge University Press, 223–243.

Wiggins, David. 1980. Weakness of Will, Commensurability, and the Objects of Deliberation and Desire. In Amelie O. Rorty (ed.), *Essays on Aristotle's Ethics*. Berkeley: University of California Press, 241–266.

—1987*a*. Truth, Invention and the Meaning of Life. In D. Wiggins, *Needs, Values, and Truth: Essays in the Philosophy of Value*. Oxford: Blackwell, 87–138.

— 1987b. Truth, and Truth as Predicated of Moral Judgments. In D. Wiggins, *Needs, Values, and Truth: Essays in the Philosophy of Value*. Oxford: Blackwell, 139–184.

— 1987c. A Sensible Subjectivism? In D. Wiggins, *Needs, Values, and Truth: Essays in the Philosophy of Value*. Oxford: Blackwell, 185–214.

— 1990–91. Moral Cognitivism, Moral Relativism and Motivating Moral Beliefs. *Proceedings of the Aristotelian Society* 91: 61–85.

— 1996. Objective and Subjective in Ethics. In Brad Hooker (ed.), *Truth in Ethics*. Oxford: Blackwell, 1–18.

— 1997. Incommensurability: Four Proposals. In Ruth Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 52–66.

— 2005. Objectivity in Ethics; Two Difficulties, Two Responses. Ratio 18: 1–26.

Williams, Bernard. 1973*a*. Ethical Consistency. In Bernard Williams, *Problems of the Self*. Cambridge: Cambridge University Press, 166–186.

— 1973b. Consistency and Realism. In Bernard Williams, *Problems of the Self.* Cambridge: Cambridge University Press, 187–206.

— 1973c. A Critique of Utilitarianism. In J. J. C. Smart and Bernard Williams, *Utilitarianism: For and Against*. Cambridge: Cambridge University Press, 73–150.

— 1981. Conflicts of Values. In Bernard Williams, *Moral Luck*. Cambridge: Cambridge University Press, 71–82.

— 1985*a. Ethics and the Limits of Philosophy.* Cambridge, Mass.: Harvard University Press.

— 1985b. Ethics and the Fabric of the World. In Ted Honderich (ed.), *Morality and Objectivity*. London: Routledge & Kegan Paul, 203–214.

— 2001. Liberalism and Loss. In Mark Lilla, Ronald Dworkin and Robert Silvers (eds.), *The Legacy of Isaiah Berlin*. New York: New York Review of Books, 91–103.

- Williamson, Timothy. 1992. Vagueness and Ignorance. Proceedings of the Aristotelian Society, Suppl. Vol. 66: 145–162.
- Winch, Peter. 1972. The Universalizability of Moral Judgments. In Peter Winch, *Ethics and Action*. London: Routledge & Kegan Paul, 151–170.
- Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Trans. G. E. Anscombe. New York: Macmillan.

Wolf, Susan. 1992. Two Levels of Pluralism. Ethics 102: 785-798.

- Wong, David B. 1984. Moral Relativity. Berkeley: University of California Press.
- Wright, Crispin. 1988. Moral Values, Projection, and Secondary Qualities. *Proceedings* of the Aristotelian Society, Suppl. Vol. 62: 1–26.

— 1992. Truth and Objectivity. Cambridge, MA: Harvard University Press.

— 1993. Realism: The Contemporary Debate – W(h)ither Now? In John Haldane and Crispin Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press, 63–84.

—— 1994. Realism: Pure and Simple? *International Journal of Philosophical Studies* 2: 327–341.

— 2003. Realism, Anti-Realism, Irrealism, Quasi-Realism. In Crispin Wright, *Saving the Differences: Essays on Themes from "Truth and Objectivity"*. Cambridge, Mass.: Harvard University Press, 11–47.

Zangwill, Nick. 1999. Dilemmas and Moral Realism. Utilitas 11: 71–90.

— 2006. Moral Epistemology and the Because Constraint. In James Dreier (ed.), *Contemporary Debates in Moral Theory*. Oxford: Blackwell, 263–281.