TRUTH AND MORAL DISCOURSE

by

FRITZ J. MCDONALD

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

2006

UMI Number: 3213155

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the degree requirement for the degree of Doctor of Philosophy.

_____	_____
Date	Chair of Examining Committee Virginia Held
_____	_____
Date	Executive Officer John Greenwood

Paul Horwich

_____

Michael Devitt

_____

Steven Ross

_____

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

TRUTH AND MORAL DISCOURSE

by

Fritz J. McDonald

Advisor: Paul Horwich

In this dissertation, I will examine the relationship between moral theory, the metaphysics of morality, and the semantics of moral discourse. In particular, I will consider what bearing the question of whether moral utterances are straightforwardly true or false has on meta-ethical positions and moral theories. I will argue that the conflation of semantic, metaphysical, psychological, and moral questions has led to a great deal of unnecessary confusion regarding many of the most critical concerns in these domains. In order to resolve this confusion, I will present an account of the semantics of moral discourse based on our commonsense view of the practice of moral argument, the minimalist theory of truth, and a use theory of meaning. In light of this account, I will consider how semantic issues do and do not bear on certain central metaphysical and moral problems.

My main conclusions are: (1) Minimalism and the use theory have the resources to present an account of the semantics of moral discourse fully consistent with common sense; (2) There are no grounds for holding that moral utterances are not truth-apt, and any theory that denies this critical claim will run into serious difficulties; (3) There is no inconsistency between minimalism and the core principles and motivations behind important meta-ethical theories such as expressivism and constructivism.

PREFACE

In this dissertation, I will examine the relationship between moral theory, the metaphysics of morality, and the semantics of moral discourse. In particular, I will consider what bearing the question of whether moral utterances are straightforwardly true or false has on meta-ethical positions and moral theories. I will argue that the conflation of semantic, metaphysical, psychological, and moral questions has led to a great deal of unnecessary confusion regarding many of the most critical concerns in these domains. In order to resolve this confusion, I will present an account of the semantics of moral discourse based on our commonsense view of the practice of moral argument, the minimalist theory of truth, and a use theory of meaning. In light of this account, I will consider how semantic issues do and do not bear on certain central metaphysical and moral problems.

My main conclusions are: (1) Minimalism and the use theory have the resources to present an account of the semantics of moral discourse fully consistent with common sense; (2) There are no grounds for holding that moral utterances are not truth-apt, and any theory that denies

this critical claim will run into serious difficulties; (3) There is no inconsistency between minimalism and the core principles and motivations behind important meta-ethical theories such as expressivism and constructivism.

The structure of the dissertation is as follows:

Chapter 1: Minimalist and Deflationary Theories of Truth

This chapter presents an account and brief defense of deflationism regarding truth and, in particular, a defense of the minimalist theory of truth. A number of the recent debates regarding the relationship between deflationary truth and meta-ethical theories have been framed with reference only to the minimalist theory. The concerns regarding the compatibility of minimalism and these meta-ethical theories can also be raised for other, competing deflationary theories of truth. I will describe these other deflationary theories (such as disquotationalism and prosententialism) with the intention of showing how the account of moral discourse presented here could generalize to resolve similar issues for these theories of truth that are similar to minimalism in many important respects.

Chapter 2: On Considering Moral Utterances True or False

I argue that the ordinary practice of moral discussion and argument gives us reason to regard moral utterances as straightforwardly true or false. I will contrast this position with other positions on the truth or falsehood of moral utterances, and argue that the competing positions cannot make sense of our ordinary practice of arguing about what one ought to do.

Chapter 3: A Use-Theoretic Semantics for Moral Terms

The third chapter of the dissertation will present an account of the semantics of moral terms in terms of the use of such terms. This chapter will draw on the work of other philosophers who have developed use theoretic and conceptual role accounts of meaning (generally) and the meaning of moral terms (specifically). Such an account of meaning, spelled out in terms of use rather than truth-conditions, is required on a minimalist theory of truth. I will argue that the use theory of meaning has significant advantages over theories of meaning that attempt to explain meaning in terms of truth or reference.

Chapter 4: Minimalist Schemas: Truth, Properties, and Facts

In this chapter, I will develop the view that minimalism regarding truth can be extended, through schemas related to the minimalist truth schema, to provide an account not only of truth, but also of facts and properties. In light of these schemas and the arguments presented in Chapter 2, there is a clear sense in which moral utterances are factual utterances, and moral terms refer to moral properties. In order to make sense of metaphysical disputes, I will explain how this account allows for a distinction between robust and nonrobust properties.

Chapter 5: Morality, Metaphysics, Semantics, and Psychology

The following three questions are among the main concerns of philosophers interested in the metaphysics of morality, moral discourse, and moral psychology:

1. Is it possible for utterances in moral discourse, such as 'Rape is wrong' and 'Great inequalities in the distribution of wealth are unjust,' to be true or false?
2. Are there moral properties, such as goodness and justice? Are such properties ones that figure into scientific or causal explanations?

3. Are moral utterances such as 'Rape is wrong' expressions of moral beliefs, or merely of non-doxastic mental states such as desires or emotions?

Prima facie, the first question is a metaphysical question, regarding the status of a certain class of properties. The second question is question regarding (in part) the semantics of truth and the practice of moral discourse, and is concerned with the truth or falsehood of certain utterances. The third question is a psychological question, concerned with the role of moral thoughts in a fully developed belief-desire psychology. One of the aims of this dissertation is to defend these *prima facie* views of these three questions against the common philosophical view that the answers to any one of these questions implies an answer to all of the others. The failure to keep the positions that answer questions 1, 2, and 3 separate has led to significant confusion regarding the relationship between metaphysics, moral psychology, and the role of moral utterances.

Chapter 6: On The Purported Incompatibility of Expressivism and Deflationism

In the sixth chapter, I will consider a problem that has been discussed extensively in recent literature on moral discourse and moral theory. The difficulty is that it does not seem possible to formulate certain moral views, including views usually termed "emotivist" or "expressivist," if one accepts a minimalist or other deflationary theory of truth. In this chapter, I will resolve this difficulty in light of the framework developed in Chapters 1-5; and I will consider and criticize a number of recent attempts to solve this problem. This will be a fairly extensive chapter, containing several subchapters.

Chapter 7: Constructivist Ethics and Deflationary Truth

In the final chapter, I will discuss a related difficulty that has received far less discussion in the recent philosophical literature. Constructivist moral theories, such as the view John Rawls details in his "Kantian Constructivism in Moral Theory," raise concerns similar to those raised by expressivism. Rawls claims that constructivists should not regard moral principles as either true or false. Rather, these principles should be considered only reasonable or unreasonable, for they do not

pertain to a moral reality independent of reasoning beings. I will consider Rawls's reasons for holding this position on truth, argue that such a position on truth is inessential to constructivism, hence there is no incompatibility between constructivism and a minimalist (specifically) and deflationist (generally) accounts of truth.

ACKNOWLEDGMENTS

During my four years of studying philosophy at Rutgers College and my five years of studying philosophy at the CUNY Graduate Center, I have had the opportunity to study with an almost absurdly large and varied group of philosophers who have given me a great deal of advice, help, and knowledge along with sometimes terribly difficult challenges. Along with the professors acknowledged below, I owe the greatest debts to Paul Boghossian, Frances Egan, Saul Kripke, Brian Loar, Ernie Lepore, Brian McLaughlin, David Rosenthal, and Chin-Liew Ten.

Two professors have played a special role in inspiring my interest in the subjects discussed in this dissertation. One of the central ideas in this dissertation, regarding the relationship between inflationary theories of truth and constructivism, was suggested in class by Thomas Nagel. I am grateful to Professor Nagel for allowing me to sit in on his NYU Ethics survey course, where I began to be greatly interested in ethics and meta-ethics. In my first semester at the Graduate Center, Jerry Katz taught the most difficult course I have taken as a student, Philosophy of Language. In the next semester, he shared with us the manuscript of his last book, *Sense, Reference, and*

Philosophy. While discussing the ideas of this book with a tough, quick mind, I realized both that I would like to write a great deal more on the issue of truth and that I may perhaps have some philosophical ability. I miss him terribly.

Anyone who reads this dissertation will realize how much I owe intellectually to the work of my dissertation advisor, Paul Horwich. To a great degree, the positions presented here in this dissertation are elaborations and defenses of positions presented by Paul in his work. I am grateful to Paul for what I have learned from him in classes, in personal discussions, and from his books, which set the bar very high for clarity and philosophical rigor.

I could not have asked for a better dissertation committee. Steve Ross has had a great deal of enthusiasm for this project since I told him that I was writing on meta-ethics and truth, and I hope that the final product lives up to some of his expectations. I am most grateful to Steve for not letting me lose sight entirely of the truly important issues regarding morality. I am grateful to Michael Devitt not only for what I have learned from his classes and writings, but also for learning from his example as a philosopher while working as his work study assistant. I am also deeply grateful to Virginia Held and

Catherine Wilson for agreeing to be examiners at my dissertation defense, and I hope to answer some of their most difficult questions in the work I do during my philosophical career.

I have received financial support as well as valuable teaching experience over the first four years of the PhD program at Hunter College and during the final year at the CUNY Law School. I am particularly grateful to James Freeman and Frank Kirkland of Hunter College for their work as mentors. In addition, I am grateful to the following organizations and schools for providing me with a forum to present my work, and (in many cases) for financial support to travel to conferences: the CUNY Graduate Center, the American Philosophical Association, the Long Island Philosophical Society, the New Jersey Regional Philosophy Association, the Society for Student Philosophers, Fordham University, Rutgers University, Princeton University, the University of California—San Diego, the University of Pittsburgh, and Carnegie Mellon University. I am also grateful to the New York Public Library—A great deal of the work on this dissertation was done at the NYPL Humanities and Social Sciences Library on 42nd Street.

I would not have been able to do any of the philosophical work that I have done over the last ten years

without the love and support of Susan J. Cronin. This dissertation is dedicated to her.

TABLE OF CONTENTS

CHAPTER 1

MINIMALIST AND DEFLATIONARY THEORIES OF TRUTH

When one says that it is true that the Earth moves, what is the purpose of the predication of 'truth'[1] to the proposition that the Earth moves? Is a property, the property of being true, attributed to this proposition? Is this a property with a substantial, underlying nature, a property in need of the same sort of detailed analysis as properties such as being an electron, being in a conscious mental state, or being a just society? Is this a property of fundamental philosophical as well as general theoretical importance? Does this property play a significant role in accounts of the nature of meaning? Questions of this kind have been of central importance in philosophical debates over the nature of truth.

According to certain theories of truth, the answer to all or nearly all of these questions (with the obvious exception of the first, rhetorical question) is yes. Theories that answer a majority of these questions in the affirmative are typically labeled "inflationary" accounts of truth. On an inflationary conception, when one asserts

---

[1] Throughout this dissertation, I will use single quotes when mentioning words and sentences. Double quotes will be used for direct quotations from other sources as well as for scare quotes.

that it is true that the Earth moves, one attributes a property to the proposition that the Earth moves, and the property thus attributed stands in need of a philosophical analysis. The account given of this property by the theory often plays a significant role in explaining other phenomena, such as the nature of meaning or the success of scientific theories.

The debates regarding the relationship between truth and meta-ethics that I will discuss throughout this dissertation are often shaped by the assumption, either implicit or explicit, that some form of inflationism about truth is true. These debates often overlook important differences among deflationary theories as well. In order to clarify these debates regarding truth and meta-ethics, I will in this chapter briefly summarize and criticize one of the most well developed and thoroughly discussed inflationary theories, the correspondence theory. In the following section, I will provide a characterization of deflationism, and then detailed accounts of five deflationary theories. The hope is that, by spelling out these theories in detail, I can avoid some of the troubles that have arisen in the philosophical literature due to the conflation of distinct theories of truth and the oversimplification resulting from assuming terms such as

'inflationism' and 'deflationism' refer to a single well-defined theory as opposed to a family of theories that have several, but not all, features in common.

Inflationism and Correspondence

Inflationary theories, as I mentioned above, involve identifying truth with a property, and then giving an analysis of the underlying nature of that property. Pragmatist theories of truth identify truth with the property of being a useful proposition to believe (James 1907); Proof theories of truth[2] identify truth with provability (Peirce 1878, Dummett 1978, Putnam 1981, Misak 2000); Coherence theories of truth identify the truth of a set of beliefs with the property of being a system of beliefs that fit together well (Blanshard 1939, Walker 1989). There are good reasons to doubt that these accounts of truth accurately capture our notion of truth, given that the set of true propositions diverges from the set of propositions that are useful, provable, or coherent[3].

[2] Proof theories of truth are also referred to as pragmatist theories, pragmaticist theories (Peirce's term), and verificationist theories. I will use the term 'proof theory' to avoid confusion with the verificationist theory of meaning and the pragmatist theory that identifies truth with utility.

[3] There is an excellent discussion of inflationary theories and their problems in Horwich 1998b, pp. 8-10.

One inflationist theory of truth, the correspondence theory of truth, will be discussed often throughout this dissertation due to its influence upon debates in meta-ethics. The correspondence theory identifies the property of being true with the property of being a proposition that corresponds to the facts[4]. This rough characterization, which I will discuss further below, obviously requires further characterization itself. Spelling out what a fact is, and what it would be for a proposition to correspond to such a thing, is a major task of the correspondence theory.

There are several well-known difficulties with the correspondence theory of truth. These difficulties will be discussed only briefly here[5]. Exactly what a fact is and what it would mean for a proposition to correspond to a fact is obscure. Is a fact some sort of sentence-shaped object in the world? Is the correspondence some kind of resemblance? What are the criteria of identity for facts?[6]

---

[4] Classic presentations of the correspondence theory of truth include Russell 1912, Wittgenstein 1921, and Austin 1950.

[5] Strawson 1950 offers influential and important criticisms of the classical correspondence theory. The crucial problems for correspondence theories spelled out in terms of causal theories of reference are discussed in Horwich 1998b, pp. 113-117.

[6] The difficulties presented for the correspondence theories by the need to spell out the nature of facts will be discussed in detail in Chapter 4, **Minimalist Schemas: Truth, Properties, and Facts.**

In order to clarify some of these obscurities, philosophers influenced by Hartry Field's classic paper "Tarski's Theory of Truth" have attempted to spell out the nature of this correspondence in terms of the reference of subsentential units[7]. These referential relations are then, in turn, explained in terms of a causal theory of reference[8]. However, the philosophers carrying out this project have run into serious difficulties. How does one assure that the causal links between a term such as 'Earth' and the Earth itself are specified in the proper way to explain reference? How does a causal theorist of reference distinguish appropriate causal chains from inappropriate causal chains? Complex theories presented by philosophers such as Fred Dretske (1988) and Jerry Fodor (1990) attempting to specify the proper causal link between words and the world have been found lacking[9].

Deflationary Theories

Philosophers who have been skeptical of the analyses given by inflationary theories such as the correspondence

[7] Whether or not Field's interpretation of Tarski is correct will be discussed in detail below in the section **Tarski's Theory.**

[8] Devitt 2001 offers an elaboration and defense of this kind of correspondence theory.

[9] Barry Loewer (1987) details the problems presented by misrepresentation for Dretske's theory, and Fred Adams and Ken Aizawa (1994) raise serious problems for Fodor's attempted resolution of the misrepresentation problem.

theory have asked whether it is a mistake to assume that there is a property of truth with a substantial underlying nature. Is it necessary to give such an account of the property of truth in order to explain the function of the predicate 'true'? Or is there a different account that fully explains the function of this predicate? If one can give a full account of the function of the truth term without appeal to one of these vexed theories of the underlying nature of truth, why would any further theorizing be required?

In the following sections of this chapter, I will summarize five well-known deflationary theories of truth: the redundancy theory, Tarski's theory of truth, the disquotational theory, the prosentential theory, and the minimalist theory. This is not intended to be a complete account of all deflationary theories, nor of all variations on these deflationary theories. There have been other deflationary theories proposed in the past that I will not discuss, such as the performative theory and the sentence-variable theory, and it would be highly surprising given the level of recent interest in deflationism if a number of other such theories are not proposed in the near future. However, these summaries of these selected theories will aim to indicate what is at issue in the debates that will

be discussed throughout this dissertation, debates over deflationary theories and their philosophical import. It should be clear, once I have spelled out the approach that I think should be taken by proponents of any of these five theories to the central issues in meta-ethics, how to extend this approach for other present or future deflationary theories.

The summaries I will present below will, I hope, make it clear that regardless of the application of the unitary label 'deflationist' to these theories, there are significant differences among these theories of truth. Each of these theories takes a different position on the role of the truth predicate. These theories differ on the issue of whether or not truth is a property. Some of these theories involve complexities such as appeals to substitutional quantification, whereas others do not.

A significant difference among theories of truth regards the matter of which entities are the bearers of truth, if indeed there are bearers of truth. Those deflationary theories that do not consider truth to be a property, such as the redundancy theory and prosententialism (discussed further below), hold that there are no bearers of truth. According to such theories, 'true' is not a genuine predicate—There are no predications of

'true' at the level of logical form. In light of this view on the function of 'true,' these theorists deny there is a property of truth. Thus it is inappropriate to talk of a "bearer of truth" to which this property is attributed. As Dorothy Grover writes regarding her prosentential account of truth, on such a theory "[t]ruth bearers and the property truth drop out…" (Grover 1995, 710).

On the other hand, if truth is a property, to what is this property properly attributed? One could hold that truth is attributed to entities such as sentences, utterances, beliefs, or propositions. A theory might hold that certain bearers of truth are basic: attributions of truth to such bearers account for all attributions of truth to other bearers. As I will discuss below in the section on minimalism, it is an important aspect of the minimalist theory that basic attributions of truth are to propositions, and the fact that truth is also a property of sentences, beliefs, and utterances is accounted for in terms of the attribution of truth to propositions. Sentences and utterances express propositions, and it is due to the fact that a sentence (or utterance) expresses a true proposition that we call such a sentence (or utterance) true. Beliefs have propositional contents, and it is in virtue of the fact that the propositional contents

of certain beliefs are true that we call such beliefs true. There are, as I will note below, significant difficulties that arise for theories that claim the basic attributions of truth are to sentences rather than propositions. To reiterate, among those deflationary theories holding that truth is attributed to entities, various positions could be taken regarding the entities that are the bearers of truth.

It is not uncommon to find objections raised against particular deflationary theories that involve ignorance of the difference between one deflationary theory and another. For instance, the minimalist theory is sometimes criticized for denying that there is a property of truth, even though Horwich (1998b) quite clearly claims that minimalism holds that there is a property of truth.

While these theories may differ in important respects, there are certain features these theories have in common that give rise to the issues I will discuss in this dissertation. As I will explain in detail in Chapter 2, all of these theories share the feature of, to varying degrees, trivializing the distinction between asserting that p and asserting that p is true. This shared feature plays a significant role in the debates over truth and its relation to meta-ethical controversies that I will focus on in Chapters 5, 6, and 7.

The Redundancy Theory of Truth

The oldest formulation of a deflationary theory of truth is the redundancy theory. Such a theory is found in the work of Gottlob Frege (1918-19), Frank Ramsey (1990), and A.J. Ayer (1936). The redundancy theorist claims that the assertion that p is synonymous with the assertion that p is true. Hence, there is no difference in meaning between the assertion that the Earth moves, and the assertion that it is true that the Earth moves. Ayer summarizes the view nicely:

> …[W]e find that in all sentences of the form "*p* is true," the phrase "is true" is logically superfluous. When, for example, one says that the proposition "Queen Anne is dead" is true, all that one is saying is that Queen Anne is dead (Ayer 1936, 88).

The predicate 'true' is, as the name of the theory suggests, merely redundant whenever it occurs in a sentence. Given that the insertion of 'true' into a sentence adds no additional content to that sentence, we could conceivably eliminate all occurrences of such a term in a language with no loss in content.

There are well-known problems with such a view. First, it would be difficult for a redundancy theorist to

account for the role of the truth term in sentences such as "What Kerry said about the Iraq war is true." It would be nonsensical to claim that "What Kerry said about the Iraq war is true" and "What Kerry said about the Iraq war" are synonymous, for the former is a complete sentence and the latter is a phrase referring to a number of assertions. Ramsey attempted to solve this problem by introducing to the theory an account of the role the truth term plays in such attributions of truth. On Ramsey's account, we are to explain the attribution of truth in "What Kerry said about the Iraq war is true" as follows:

R: For all x, if Kerry said regarding Iraq that x, then x.

The trouble with this attempted solution to this problem is, as Horwich (1998b) notes, that it requires quantification into an opaque context, and is thus ill formed unless a language contains substitutional quantifiers. Substitutional quantifers, however, are problematic for a number of reasons[10].

Second, the claim of synonymy between the assertion that p and the assertion that p is true is *prima facie* wrong. As Jerry Katz points out (2001), this account of the

[10] For a full discussion of this objection, see Horwich 1998b, pp. 32.

meaning of 'true' fails the Church translation test, for (4), not (3), is a proper translation into German of (2).

(1) Frege was very influential.

(2) 'Frege was very influential' is true.

(3) Frege war sehr beeinflußreich.

(4) 'Frege war sehr beeinflußreich' ist wahr.

This suggests that predicates such as 'true' and 'wahr,' *pace* the redundancy theorist, do add additional meaning to the sentences in which they occur.

Tarski's Theory

Alfred Tarski set a condition of adequacy for any theory of truth: Any such theory has to be materially adequate. For a theory of truth to be materially adequate, it has to entail every instance of the schema T:

> (T) X is true in L, if and only if p. (Tarski 1944, 335)

It is important to note a major restriction of a Tarskian theory—Tarskian truth definitions are not definitions of the concept of truth (generally), but of truth in a particular language. The bearers of truth, on a Tarskian theory, are sentences. In schema T, p is a placeholder for a sentence in the object language L, for which truth is

defined, and X is a quoted expression in the metalanguage[11]. Thus a materially adequate theory of truth will produce instances of schema T such as T1:

(T1) 'The Earth moves' is true in English, if and only if the Earth moves.

Tarski's theory attempts to produce a materially adequate theory of truth for a particular language by first presenting a description of the semantics of names and predicates in a language. An account is presented as well of the semantics of logical terms such as quantifiers and connectives. In light of this account, the theory shows how infinitely many sentences in a language can be built recursively from names and predicates through the use of the logical connectives and quantifiers. I will present an example of how such a recursive construction of sentences is carried out below.

The description of the semantics of names and predicates consists of a list of principles, listing names and the objects they denote, and predicates and the properties of the objects that satisfy such predicates.

---

[11] In this example, for the sake of simplicity, I am using English as both the object language and the metalanguage. The concerns regarding the semantic paradoxes that might lead one to regard using English as both the object language and the metalanguage problematic are outside of the scope of this dissertation.

Thus, for the expressions that make up the sentence contained in schema instance T1, we would appeal to principles such as D and S:

D: the expression 'the Earth' designates the Earth.

S: things that move satisfy the sentential function 'x moves'

In order to explain how other sentences are built up from their parts we appeal to principles in the theory spelling out the role played by logical connectives, and quantifiers.

For instance, to explain the construction of the sentence 'The Earth does not move,' we would appeal to D, S, and principle C:

C: ~p is true if and only if p is not true.

Appealing to principles D, S, and C, one can derive another instance of the T schema, T2:

T2: 'The Earth does not move' is true in English if and only if it is not true that the Earth moves.

By repeated appeal to the principles specifying the role played by names, predicates, and connectives, it is easy to see how such a theory could generate infinitely many instances of the T schema.

The issue of whether the best formulation of Tarski's theory of truth is as a deflationary or inflationary theory

of truth has been disputed by Hartry Field (1972) and Scott Soames (1984). Field contends that the Tarskian theory is incomplete unless principles regarding denotation such as D and principles regarding satisfaction such as S are explained by appeal to a naturalistic account of these relations, such as a causal theory of reference. As noted above in the section **Inflationism and Correspondence**, contemporary philosophers attempting to defend a correspondence theory of truth have taken such an approach, relying on the work of Tarski as well as the theory of reference developed first by Kripke (1980). In his response to Field, Soames contends that the definition of satisfaction and denotation in terms of lists of principles such as P1 and P2 is adequate, and regarding such lists as adequate is important to understand the "deflationist character" of Tarski's view (Soames 1984, 414). Once truth has been reduced to the primitive relations of denotation and satisfaction, no further theorizing or explanation is required in order to fully understand the notion of truth.

The issue of the proper way to pursue Tarski's program is beyond the scope of this dissertation. It is important to note, however, that there are serious difficulties for a Tarskian theory of truth, whether deflationary or inflationary, difficulties that do not plague deflationary

approaches such as the disquotational theory, prosentential theory, and minimalism. It is clear, as Tarski himself recognized, that we cannot account for the notion of truth for natural languages such as English using a Tarskian theory. Tarski had doubts regarding the possibility of presenting an account of the truth for a natural language such as English due to the liar paradoxes: Using the resources of such languages, we can form paradoxical sentences such as 'This sentence is not true.' The most significant problem in this regard is how to properly account for the truth of sentences such as subjunctive conditionals, modal statements, and propositional attitude attributions by appeal to the resources used to construct sentences on a Tarskian account. Until a satisfactory theory of such constructions is provided by Tarskians, such a theory of truth is incomplete.

The Disquotational Theory of Truth

Unlike the Tarskian account of truth, the disquotational theory of truth proposed by Quine (1970, 1992), Leeds (1978), and Field (1986) does not require that truth be accounted for in terms of satisfaction, denotation, and recursive rules for sentence construction. Rather, this account claims, as Quine puts it, that "truth is disquotation" (Quine 1992, 80). As was the case

according to the Tarskian theory, the bearers of truth on a disquotational account are sentences. In fact, to properly spell out the theory, it must be made clear that truth is only attributed to a restricted set of sentences: For the eternal sentences of a language that is understood by a speaker[12], the speaker will recognize that 'p' is true iff p. The speaker of the language can appeal to truth in order to disquote the sentence mentioned on the left-hand side of this biconditional. This fact, also noted by the Tarskian theory, is regarded as basic on the disquotational account, requiring no further explanation. Anyone who understands English and has a grasp of the notion of truth will recognize that D1 and similar instances of the disquotational schema are acceptable:

D1: 'The Earth moves' is true iff the Earth moves.

As noted above, the disquotational theory, unlike the minimalist account, does not appeal to propositions as the bearers of truth. The bearers of truth on the disquotational account are a restricted class of sentences, the eternal sentences. Eternal sentences are context-independent sentences. This requirement is a significant one for this account. For instance, it would be

[12] The reasons for the restriction of sentences to which truth is attributed to only those sentences understood by a speaker and to eternal sentences will be explained below.

troublesome if the left side of the following schema instance concerned a sentence uttered on one context (say, July 21, 2004) and biconditional itself is uttered in a different context (July 22, 2004):

D2: 'It is Tuesday today' is true iff it is Tuesday today.

Thus the only proper candidate sentences for instances of the disquotational schema are eternal sentences, sentences with truth values not dependent upon context, such as 'July 20, 2004 is a Tuesday.' The attribution of truth to sentences only, and not to propositions, is appealing to philosophers who are dubious of the existence of propositions.

Unlike the redundancy theory, the disquotational account does not assume that the notion of truth plays no significant role in the language. On this account, it is correctly noted that truth plays the important role of allowing one to formulate generalizations about true sentences. For instance, it would be impossible for a redundancy theorist to account for the fact that all sentences with the logical form p∨~p are true. One can only assert particular instances of this schema, p∨~p, and attribution of truth to particular instances of this schema are eliminable redundancies. The disquotational account

allows one to semantically ascend from each instance of the schema p∨~p to the metalinguistic level. Take particular instances such as: 'the Yankees will win the World Series ∨ the Yankees will not win the World Series'; and 'the Red Sox will win the World Series ∨ the Red Sox will not win the World Series.' We can then ascend, via the disquotational schema, to the metalinguistic level to assert that 'The Yankees will win the World Series ∨ the Yankees will lose the World Series' is true, along with all of the other instances of this schema. Such semantic ascent allows one to assert that the conjunction of all instances of this schema p∨~p are true. By allowing for the construction of such infinite conjunctions, the disquotational account explains the important role played by attributions of truth.

The disquotational account only explains one class of attributions of truth, namely attribution of truth to sentences. How are we to explain other attributions of truth, such as attribution of truth to beliefs? One possible way to do so is to claim that truth is attributed to the propositional contents of these beliefs. However, if a disquotational theorist appeals to propositions, then

there is no significant difference between this account and the minimalist theory discussed below.

A merit of minimalism that is not shared by the disquotational theory is that the disquotational theory cannot explain how we can apply the notion of truth to sentences that we do not understand. A speaker can only comprehend instances of the schema spelled out in her own language; A monolingual English speaker would not know why it is that 'Schnee ist weiss' is true iff *schnee ist weiss*. For this reason, Field restricts the theory of truth to a specific set of utterances, "only…utterances a person understands" (Field 1994, 405). This limitation in the ability of the disquotational theory to explain the concept of truth—limiting the concept to one that only applies to the utterances one understands—is a consequence of the disquotational theorist's refusal to countenance propositions. Without such restrictions, other theories such as minimalism can avoid this limitation of the disquotational theory.

The Prosentential Theory of Truth

The prosentential theorist, like the disquotational theorist, provides an account of the role of the truth predicate in a language. The most significant difference between the prosentential theory and all of the other

deflationary theories of truth is the distinctive account the prosentential theorist gives of the role played by the truth predicate. The prosentential theory claims that assertions such as 'That is true' have a function analogous to pronouns.

On one reading of sentence AP, the pronoun 'he' is an anaphoric pronoun:

AP: Derek knew that he needed to hit a home run.

The pronoun 'he' has the same referent as its antecedent, the name 'Derek.' It obtains this referent by being anaphorically dependent upon the antecedent. In addition to anaphoric pronouns, as Grover, Camp, and Belnap (1975) point out, there are anaphoric proadjectives, such as 'so' in the following quotation from Pope: "To make men happy and to keep them so" (Grover, Camp, and Belnap 1975, 84). The expression 'so' essentially plays the same role in this sentence as a second occurrence of 'happy' would play, describing how the men being discussed by Pope are kept. The word 'so' inherits its meaning from its antecedent, 'happy.' The following discourse, DIS1, 'That is true' is, according to the prosentential theory, a prosentence:

DIS1: Galileo: The Earth moves. Castelli: That is true.

The prosentence ‘That is true,’ asserted by Castelli, is anaphorically dependent upon its antecedent, Galileo’s assertion ‘The Earth moves.’ Just as the pronoun ‘he’ in AP inherits its content from its antecedent and the proadjective ‘so’ in the Pope quotation inherits its content from the adjective ‘happy,’ according to prosententialism the prosentence ‘That is true’ has the same content as its antecedent. Thus, in this context, ‘That is true’ means that the Earth moves.

The prosentential theory has to contend with one of the difficulties that plagued the redundancy theory of truth. The bare-bones prosentential theory summarized above does not have the resources to explain the meaning of sentences such as K1:

K1: K1: Everything Kerry said about the Iraq war is true.

In order to explain such occurrences of the truth term, Grover, Camp, and Belnap (1975) claim that the English sentence K1 is equivalent to the sentence K2:

K2: Everything is such that if Kerry said regarding the Iraq war that it is true then it is true.

There are two things to note about this explanation of the meaning of K1. First, it requires reading the pronoun ‘it’ contained within ‘it is true’ not as a variable ranging

over propositions, but as a syncategorematic part of the anaphoric expression 'it is true.' This expression has its content due to its anaphoric dependence upon what John Kerry said regarding Iraq. This complex account of the meaning of K1 seems a bit *ad hoc*, required only to maintain the prosententialist claim that true is not a genuine predicate. A more plausible account of equivalence of K1 and K2 would claim that 'true' (*pace* prosententialism) is a genuine predicate, and the pronoun 'it' is a variable ranging over propositions, as minimalism suggests[13].

The Minimalist Theory of Truth

The minimalist theory claims regarding the meaning of the term 'true' in English (and the meanings of similar terms in other languages) that it is best analyzed in terms of a fact regarding the use of the term by speakers of the English language. The meaning of 'true' is explained fundamentally by the acceptance of a trivial schema T:

T: <p> is true iff p.

In the schema, '<p>' is short for 'the proposition that p.'

Speakers of English are inclined to accept, for any given proposition, <p>, that the proposition that p is true iff p. According to minimalism, the fact that speakers

[13] For an account of how minimalism handles such expressions, see Horwich 1998b, 31-33.

accept instances of such a schema explains the purpose of the notion of truth, which is to allow one to form generalizations such as 'Everything the president said in his speech was true' and 'All instances of 'if p, then p' are true.' The generalizing role of truth is the sole purpose of the notion of truth. No further facts, beyond acceptance of the schema, are required in order to specify the meaning of the term 'true.'

The minimalist theory offers an account not only of the meaning of 'true' and the purpose of the notion of truth, but also an account of the property of truth. The minimalist theory, unlike deflationary theories that do not consider 'true' to be a genuine predicate (such as the redundancy theory and prosententialism), claims that truth is a property. As I have noted above, serious difficulties arise for those deflationary theories that claim that 'true' is not a genuine predicate.

Truth is, on the minimalist theory, a property of a specific sort—Unlike substantive properties such as being an electron or being a dog, the property of being true is not a property that has an underlying nature, a nature that requires a theoretical analysis. It is a mistake, Horwich (1998b) argues, to assume to assume that 'is true' must, as 'is an electron' does, attribute a substantive property.

To do so would be a mistake of the kind Wittgenstein (1953) noted as characteristic of philosophical error: One would be making the mistake of assuming on the basis of the analogous structures of 'is an electron' and 'is true' that these expressions serve the same function. Once it is recognized that it is not the function of 'is true' to refer to a substantive property, and that the minimalist account of the meaning of 'true' and the role played by truth attributions is all that needs to be said in order to explain what truth is, the confusion and error that have resulted from assuming truth must have an underlying nature are resolved.

A further important aspect of minimalism that I have discussed previously (in **Deflationary Theories**) is the account given by minimalism of the bearers of the property of being true. On the minimalist account, propositions are the bearers of truth, and all other attributions of truth are explained on the basis of attributions of truth to propositions[14]. Considering attributions of truth to propositions to be basic has significant advantages over other approaches, for (as I argued in **The Disquotational Theory of Truth**) taking attributions of truth to other

[14] For a full account of explaining attributions of truth to utterances and beliefs, see Horwich 1998b, 133-34.

entities such as sentences to be basic will leave other attributions of truth unexplained.

For this reason and others that I have discussed above in the sections on the competing deflationary theories of truth, minimalism has a number of advantages over its competitors. Regardless of these differences, as I have noted, there is one key similarity between all deflationary theories. Each deflationary theory makes trivial the distinction between asserting that p and asserting that it is true that p. Whether or not this trivialization of this distinction is troubling will be considered in the next chapter.

CHAPTER 2

ON CONSIDERING MORAL UTTERANCES TRUE OR FALSE

An extensively discussed question in the philosophical field of meta-ethics is whether utterances pertaining to normative matters generally and moral matters specifically are capable of being either straightforwardly true or straightforwardly false. Normative and moral utterances, such as utterances of 'Rape is wrong,' 'Great inequalities in the distribution of wealth are unjust,' and 'One should be polite in the company of strangers' are often claimed to be different from nonnormative and nonmoral utterances: Normative and moral utterances are purportedly incapable of being true or false. Philosophers have held the view that such utterances are incapable of being true or false for a number of reasons: Such utterances do not involve the attribution of properties or a statement of facts; Such utterances do not involve the expression of beliefs, but rather only the expression of desires; Such utterances are employed in practical rather than theoretical reasoning, and play a role in planning and action rather than describing the world. In light of the view that moral and normative matters exhibit one or a number of these

differences[15], certain philosophers have denied that normative and moral utterances are capable of being true or false.

The various views that could be held regarding truth and moral discourse are as follows: One could hold that such utterances are either true or false, and therefore in this respect do not differ from the nonmoral and nonnormative utterances. One could hold that such utterances are neither true nor false, and thus differ from the nonmoral and nonnormative utterances in this respect. Yet another view that marks a distinction between the moral/normative and the nonmoral/nonnormative in terms of truth would be a view that holds that such sentences are capable of being true or false, but only in some distinctive way that indicates the difference between normative/moral utterances and nonnormative/nonmoral utterances.

---

[15] I am raising these issues briefly here in order to indicate why certain philosophers have claimed moral utterances are incapable of being true or false. In order not to complicate the discussion at this early stage, I have presented simplified characterizations of these concerns. The issue of the nature of moral meaning and concepts will be discussed in detail in Chapter 3, the issue of moral facts and properties will be discussed in Chapter 4, the relation of these issues to moral psychology will be discussed in Chapter 5, and the bearing of such issues on expressivist theories and constructivist theories will be discussed in Chapters 6 and 7.

In order to simplify the subsequent discussion of this issue in the dissertation and to avoid repetition, I will use the following terms to refer to the theses discussed in the previous paragraph. I will call the approach that allows for the straightforward attribution of truth and falsehood to normative and moral utterances the Attribution Thesis. The view denying that truth and falsehood can be attributed to normative and moral utterances will henceforth be called the Denial Thesis. Finally, the theories calling for truth and falsehood of a distinct kind to be attributed to normative and moral utterances will be called instances of the Anomaly Thesis. In the following sections, I will discuss the motivations, based in ordinary practice and the deflationary theories of truth, for holding the Attribution Thesis. In the following sections, I will consider the prospects for the Denial and Anomaly Theses.

Moral Truth and Ordinary Practice

The claim that normative and moral utterances are incapable of being true or false does not fit well with ordinary practice. Speakers commonly attribute truth and falsehood to moral utterances. People generally consider utterances of 'Rape is wrong' true, and consider utterances of 'It is permissible to attack strangers for fun' false.

Any philosophical qualms regarding the attribution of truth to moral utterances do not seem to affect this practice of attributing truth to moral utterances.

Normative and moral assertions also have all of the same surface features as nonnormative and nonmoral assertions. 'Killing is wrong' appears to attribute wrongness to killing in just the same way that 'The Earth is round' attributes roundness to the Earth. Taking this surface structure into consideration provides a *prima facie* reason for regarding moral utterances as similar in other respects to nonmoral utterances.

In addition to sharing surface features with nonmoral discourse, attribution of truth to moral utterances is required to account for the role such utterances play in arguments. Take an argument such as:

> P1: If murdering innocent people is always wrong, then murdering a small group of innocent people to save the lives of a larger number of innocent people is wrong.
>
> P2: Murdering innocent people is always wrong.
>
> C: Murdering a small group of innocent people to save the lives of a larger number of innocent people is wrong.

Such an argument certainly appears to have the form of a valid argument, an instance of *modus ponens*. However, if

we reject the claim that moral utterances are capable of being true, one could not appeal to the truth of these claims and the form of the argument to explain why the truth of these premises would lead, necessarily, to the truth of the conclusion. Denying that truth could be attributed to moral utterances would be tantamount to claiming that there is no possibility of valid moral argument[16].

---

[16] On a prescriptivist theory of morality, such as R.M. Hare's (1997) theory, moral utterances are disguised imperatives. When one utters a sentence such as 'Rape is wrong,' one is not making a descriptive statement, but rather issuing a command. In this particular case, 'Rape is wrong' serves the purpose of making the command 'Do not rape.' If such a view were correct, one might think that moral utterances are incapable of being true or false, just as imperatives generally are incapable of being true or false. Hare develops a logic for imperatives, explaining how imperatives relate logically in a way that is analogous to the way other statements relate. Virginia Held (at the defense of this dissertation) raised the question of whether Hare's account of imperatives and logic allows for there to be valid moral arguments even if moral utterances are incapable of being true or false. I prefer to use the term 'valid' as it has traditionally been used in logic textbooks, in the sense described above in the chapter. Perhaps, if Hare's account were correct, we could use 'valid' in a looser sense, to include both arguments valid in this traditional sense as well as those arguments that are correct according to Hare's logic. While a full discussion of prescriptivism is outside of the scope of this dissertation, I find the analysis of moral utterances and the logical relations among them offered by Hare unconvincing. The main reason for this is that I do not think that prescriptivism offers an adequate account of moral motivation. (The importance of motivation for accounts of the meaning of moral terms will be discussed later, in Chapter 3 of the dissertation). Yet another

An analogous difficulty is that in order to have a notion of moral knowledge that accords well with our ordinary practice, we need to attribute truth to moral utterances. As Matthew Chrisman notes in his review of Allan Gibbard's *Thinking How to Live*, 'know' is a factive verb—One cannot claim to know a proposition unless that proposition is true. If one says

> …that normative sentences are neither true nor false, so if that's true, we cannot have normative knowledge. Yet in a Moorean vein, one might reasonably think: "Whatever I may or may not know about the semantics of normative language, I damn well know that torturing children is wrong." (Chrisman 2005, 408).

Thus the claim that moral utterances are incapable of being true creates a tension with our commonsense conception of moral knowledge.

The Quick Route to Attribution

The deflationary theories of truth discussed in the previous chapter fit well with the ordinary practice of attributing truth to moral utterances, for on a deflationary theory of truth, it would seem that there is a fairly quick route from the fact that people make sincere

---

reason is provided by the other arguments in this chapter for considering moral utterances capable of being straightforwardly true or false.

moral assertions to the claim that moral utterances are capable of being true or false. Using the resources of any of these five theories, one can show that the inference from an assertion that p to the assertion that the proposition that p is true is a trivial one. Thus as soon as one commits oneself to holding that rape is wrong, one would also commit oneself to hold (if one has a grasp of the notion of truth) that the proposition that rape is wrong is true.

In this chapter, I will show that there is such a quick route on any deflationary theory of truth, and I will offer several arguments for regarding this as a merit of these deflationary theories rather than a defect. It is safe to say that the majority of philosophers who have written on this issue assume that the quick route from assertion and deflationism to the Attribution Thesis is a problem, because it is assumed that the only way to mark the distinction between the moral and normative, on the one hand, and the nonmoral and nonnormative, on the other, is by making a distinction in terms of truth. In response to these concerns, I will discuss important distinctions between the normative/moral and nonnormative/nonmoral in the subsequent discussion of moral semantics in Chapter 3 as well as the discussion of moral metaphysics in Chapter

4. Specific considerations related to the relationship between these matters concerning the attribution of truth to moral utterances and theories such as expressivism and constructivism are too complex to discuss in detail in this chapter. I will discuss these specific theories in Chapters 6 and 7, where I indicate how attempts to evade the quick route have led philosophers to propose a number of unappealing alternative accounts of moral and normative truth.

The Quick Route: The Redundancy Theory

Take, on the redundancy theory, any sincere assertion of a moral claim, such as 'Rape is wrong.' It is quite clear that one could not distinguish this assertion from another that, according to the theory, has equivalent content: 'It is true that rape is wrong.' Thus, if the redundancy theory is correct, we cannot make any distinction at all, not even one of meaning, between a moral assertion and the assertion that a moral assertion is true.

The Quick Route: Tarki's Theory

On a deflationist Tarskian theory, we are to derive instances of the T schema from a list of principles regarding the function of connectives, how referring terms denote, and how predicates are satisfied. It is plausible

to think that a word such as ‘rape’ refers to acts of rape, thus one would expect the following to be among the list of principles spelling out denotation:

DR: The term ‘rape’ denotes rape.

We use the word ‘wrong’ as a predicate to describe certain acts, therefore the list of principles spelling out predicate satisfaction should include:

SW: Things that are wrong satisfy the sentential function ‘x is wrong.’

If DK and SW are among the principles of a Tarskian theory used to derive T schema instances, we can derive the following instance:

TR: ‘Rape is wrong’ is true if and only if rape is wrong.

Having established this equivalence, we can show that whenever one asserts that rape is wrong, one can also appeal to the Tarskian theory to infer that it is true that rape is wrong.

The Quick Route: The Disquotational Theory

The disquotational theory, unlike the Tarskian theory, makes no appeal to principles regarding predicate satisfaction, denotation, and the role of connectives. It does, however, regard T schema instances along the lines of TR as basic:

TR: ‘Rape is wrong’ is true if and only if rape is wrong.

Thus, as was the case with the Tarskian theory, this instance of the disquotational schema can be applied to a sincere assertion of ‘Rape is wrong’ to show that truth ought to be attributed by that person to the sentence ‘Rape is wrong.’

The Quick Route: The Prosentential Theory

For the prosentential theorist, the quick route from sincere moral assertion to the Attribution Thesis is illustrated by the fact that purported prosentences can be and often are used in contexts where moral assertions are the antecedents of such prosentences. Thus if Larry asserts that “Great inequalities in the distribution of wealth are unjust,” and Barry responds “That is true,” what does Barry’s assertion mean? On the prosentential account, as noted in Chapter 1, Barry’s utterance is anaphorically dependent upon Larry’s utterance, and thus his assertion ‘That is true’ has the same meaning as ‘Great inequalities in the distribution of wealth are unjust.’ There is no reason to think that we cannot use the resources of the prosentential theory to form prosentences anaphorically dependent upon moral and normative utterances in just the same way we use these resources to form prosentences

anaphorically dependent upon nonmoral and nonnormative utterances.

The Quick Route: The Minimalist Theory

On the minimalist theory, unlike some theories such as disquotationalism and prosententialism, truth is attributed not directly to sentences, but rather to propositions. On the assumption that an utterance such as 'Rape is wrong' expresses a mental state with propositional content, we would take the utterance to express the proposition that rape is wrong. Is there a reason to reject the claim that the mental state expressed in this situation does express such a proposition?

If this utterance does express the proposition that rape is wrong, then the following would be an instance of the minimalist truth schema:

TR: <Rape is wrong> is true iff rape is wrong.

Thus, any speaker with an understanding of the notion of truth, according to the minimalist theory, would be able to recognize that the claim that it is true that rape is wrong is a consequence of the truth schema and the wrongness of rape. Thus, on the minimalist theory as well as the four other deflationary theories discussed above, there is a quick route from asserting that p to asserting that p is true.

Should We Avoid the Quick Route?

Ordinary practice, as well as the deflationary theories of truth considered in chapter 1, would seem to suggest that truth can and ought be attributed to moral utterances. If a deflationist wanted to hold that truth ought not to be attributed to normative and moral utterances, such a theorist could claim that such utterances are not legitimate instances of the deflationist schemas. For example, as noted above, in order to establish that there is a quick route from sincere assertions to the Attribution Thesis on the minimalist theory, we first have to ask whether we should regard moral utterances as assertions that involve the expression of propositions. A minimalist who wanted to hold to the idea that normative and moral utterances are incapable of being true or false could deny that such utterances express propositions. Similar moves could be made on any of the other deflationary theories: A disquotationalist could deny that normative/moral utterances are legitimate instances of the T schema; A prosententialist could deny that expressions such as 'that is true' form legitimate prosentences when they have normative/moral utterances as antecedents; A Tarskian could claim that normative/moral predicates and terms fail to form sentences that are

instances of the T-schema; A redundancy theorist could deny that moral utterances are meaningful, thus rendering 'Rape is wrong' and 'It is true that rape is wrong' both meaningless.

These *ad hoc* moves may allow us to hold to the view that normative and moral utterances are incapable of being true or false, but how convincing are these views on their own merits? To reiterate a point made earlier in this chapter, any deflationist who denied that normative/moral utterances are capable of being true or false would make a claim that does not cohere with our ordinary practice of attributing truth to moral statements. Furthermore, as I also have discussed previously, denying the claim that moral utterances are capable of being true or false precludes the possibility of valid moral argument and moral knowledge.

It also runs counter to common sense to claim, as these deflationist views that avoid the quick route require us to do, that moral utterances are meaningless. Utterances such as 'rape is wrong' seem no less meaningful than utterances such as 'grass is green,' and moral/normative utterances can be embedded in all of the same contexts as any clearly meaningful, proposition-expressing utterances. For this reason, there seems to be

little motivation for a deflationist, such as the minimalist, to make the claim that is required to block the attribution of truth to moral utterances: the clear meaningfulness of moral utterances runs counter to any attempt to deny that such utterances express propositions. There seems to be neither good reason nor sufficient room to avoid the quick route.

Anomaly Theses

Perhaps philosophers who do not deny that moral utterances are capapble of being true or false may still seek to find some distinctive way of attributing a distinct kind of truth to moral utterances. One way to spell out the option of offering a distinctive kind of truth attribution for these utterances would be to hold a relativist view. Such a view holds that normative and moral utterances are true or false only relative to a particular individual or social perspective. There are good, well-known reasons to reject any such view of moral truth. Relativism does not offer us what we want when we want to condemn an act as wrong: We want not merely to say that the killing of innocents is wrong from our perspective, or relative to the prejudices of our society. We need to employ an unrelativized notion of truth to express our sincere moral commitments.

Another way to spell a theory according to which a distinctive kind of truth is attributed to moral utterances would be to claim that the theory of truth for normative and moral utterances differs from the theory one would give for other utterances, as Wright (1992) claims. Specifically, on Wright's approach to truth, there can be "a *plurality* of truth properties, qualifying as such by satisfying certain general principles…" (Wright 2001, 773). For moral discourse, Wright suggests that the property of truth that is attributed to moral utterances is what he calls 'superassertibility.' A statement is superassertible

> if and only if it is, or can be, warranted and some warrant for it would survive arbitrarily close scrutiny of its pedigree and arbitrarily extensive increments to, or other forms of improvement of, our information (Wright, ibid., 771).

The most important thing to note, for the purposes of this chapter, is that Wright contends that, while moral truth should be accounted for in terms of superassertibility, we need not claim that this is also the property of truth attributed to claims in other areas of discourse.

One reason to be dubious of the prospects for a pluralist conception of truth such as Wright's is that a pluralist theory, unlike the minimalist theory and other

deflationist views, does not give us an account of the role of the concept of truth in language[17]. Each deflationist view has a different account of the role of the concept of truth—In this discussion of pluralism, I will focus on the theory of the role of truth given by the view I consider the most plausible deflationist theory, minimalism. The minimalist theory explains the role played by truth as a device of generalization—As I discussed in Chapter 1, the minimalist theory explains how the notion of truth is employed to allow us to say things such as 'Everything the Prime Minister said was true,' or 'Every claim the author made in that book is true.'

When we make such generalized claims regarding the truth of a number of propositions, it is quite clear that we are attributing truth to propositions regarding moral matters as well as nonmoral matters. So if among the utterances of the Prime Minister are 'The killings are abhorrent' as well as 'The meeting is cancelled,' an utterance of 'Everything the Prime Minister said was true' involves the attribution of truth to the following two propositions (among others): <The meeting is cancelled>, <The killings are abhorrent>. So the role of truth as a device of generalization involves the attribution of truth

[17] I owe this argument to Paul Horwich.

to propositions expressed by moral utterances as well as nonmoral utterances.

How is the pluralist to explain the use of the notion of truth as a device of generalization? The best explanation we have of the use of truth as a device of generalization is the one offered by the minimalist theory. It would be open to the pluralist to take advantage of the resources of the minimalist theory, but the pluralist would have to do so at the cost of rejecting one of the central tenets of her view, the claim that distinct kinds of truth are to be attributed to utterances in different regions of discourse. Once again, it seems clear that there is good reason for philosophers to accept the Attribution Thesis.

CHAPTER 3

A USE-THEORETICAL SEMANTICS FOR MORAL TERMS

A clear consequence of the deflationary theories of truth discussed in Chapter 1 is that one cannot accept an account of meaning presented in terms of truth. Deflationary theories such as disquotationalism and minimalism hold that truth does not play a substantial theoretical or explanatory role, hence one cannot appeal to truth in order to explain meaning.

Furthermore, as Dummett (1959) noted, deflationary theories cannot account for knowledge of the meaning of utterances in terms of knowledge of truth conditions. This is due to the fact that, according to deflationary accounts, understanding of biconditionals such as CM constitutes our understanding of the notion of truth:

> CM: 'The cat is on the mat' is true if and only if the cat is the on mat.

Given that our understanding of such instances of a truth schema (as discussed in Chapter 1) accounts on a deflationary view for our understanding of truth, understanding of such biconditionals cannot also account for our understanding of meaning. We would need, in order

to explain knowledge of meaning in terms of knowledge of biconditionals such as CM, a prior understanding of truth that is not itself explained by appeal to understanding of the very same biconditionals. As Dummett writes:

> …[I]n order that one should gain from the explanation that P is true in such-and-such circumstances an understanding of the sense of P, he must already know what it means to say of *P* that it is true (Dummett 1959, 7).

Deflationism precludes any account of knowledge of meaning cast in terms of knowledge of truth conditions.

For these reasons, deflationists are required to reject the wide-spread view based on the work of Donald Davidson (1984), the view that meaning can be accounted for in terms of truth conditions. In his book Meaning, Paul Horwich elaborates and defends an alternative to truth conditional accounts, compatible with minimalism (and other deflationary views) regarding truth: a use theory of meaning. In this chapter, I will discuss the importance of the use theory of meaning for understanding the meaning of terms in moral discourse.

Not only is the use theory of meaning capable of accounting for the meaning of moral terms—Unlike its competitors, the use theory of meaning does not preclude

the formulation of any particular meta-ethical theories. A further important consequence of accepting a use theory of meaning is that it leads to a rejection of the alleged contrast[18] between truth-conditional discourse and non-truth-conditional discourse (Blackburn 1984). I will explain, in terms of a use theory of the meaning of moral terms, why such a contrast is mistaken.

In the first section of this chapter, I will briefly summarize the use theory of meaning. In the second section, I will present a sketch of a use theory of meaning for moral terms, drawing on the work of Sellars (1954), Harman (1987), Brandom (1994), Wedgwood (2001), and Gibbard (2003b). The aim here will not be to present a full semantics for all moral terms, but simply to sketch the basic elements of such a theory as it has been developed by these philosophers. The essence of this view, as it is nicely put by Wedgwood, is that "moral terms get their meaning from their role…in *practical reasoning*" (Wedgwood

[18] I do not intend to suggest here that there is no contrast between moral and nonmoral matters and discourse. Rather, a use theory allows for us to explain meaning on a single theory, rather than appealing, as Blackburn does, to one account of meaning for moral terms and another account for nonmoral terms. In this chapter, I draw on a use-theoretical account of the meaning of moral terms in order to distinguish moral discourse from nonmoral discourse. In Chapter 4, I develop an account of properties and facts that will allow for a distinction between so-called "realist" and "anti-realist" accounts of morality.

2001, 12). In the third section of this chapter, I will compare the prospects of such a theory of the meaning of moral terms with the prospects of explaining these meanings in theories other than the use theory. Prominent theories of meaning, theories that attempt to account for meaning in terms of reference, do not have the resources to explain the meaning of moral terms. Moreover, such theories would seem to preclude any meta-ethical theory other than a narrow range of implausible theories. The use theory, on the other hand, does not preclude the formulation of any so-called "realist" or "anti-realist" view, including the expressivist theories I will discuss in Chapter 6 and the constructivist theories I will discuss in Chapter 7.

The Use Theory of Meaning

There have been a number of philosophers who have presented theories of meaning inspired by Wittgenstein's (1953) claim that meaning is use (Sellars 1954, Field 1977, Harman 1982, Harman 1987, Block 1986, Peacocke 1992, Brandom 1994). In this chapter, I will base my discussion on Paul Horwich's use theory of meaning as presented in his book *Meaning*, an account which answers some of the most significant challenges to a use theory of meaning, such as the charge of Fodor and Lepore (1991) that use theories of meaning inevitably lead to an implausible, radical meaning

holism. This account also has the merit of not assuming that meaning is an intrinsically normative notion (*pace* Sellars 1954 and Brandom 1994), one that cannot be given a naturalistic analysis[19].

On the use theory, for any word in a language, there is a core use which constitutes the word's meaning. This use is the use of the word that provides the explanation of the word's overall use—the facts regarding the acceptance or rejection of sentences containing that word. Other factors, beyond the core use, regarding the use of the term do not figure into such explanations. This restriction to the core use avoids a radical meaning holism: "the use theory of meaning does not in fact lead to the counterintuitive form of holism" (Horwich 1998, 60), according to which "the meaning of a term depends on every single aspect of its overall use" (ibid., 61). Thus, as noted above, Horwich's theory effectively responds to the objections of Fodor and Lepore.

The core use of a word is the tendency of speakers to accept certain specified sentences containing that word in

[19] For defense of this claim, see Horwich 1998a, Chapter 8. Fodor and Lepore present a convincing argument against Brandom's claim that meaning must be explained as an intrinsically normative notion in "Brandom's Burdens" (Fodor and Lepore 2001).

certain situations. An example provided by Horwich is the core use of the word 'red':

> the explanatorily fundamental acceptance property underlying our use of the word "red" is (roughly) the disposition to apply "red" to an observed surface when and only when it is clearly red (ibid., 45).

The acceptance property of the term 'red' is the tendency to accept sentences such as 'that is red'--what Horwich calls the application of 'red'--in the presence of red objects.

For any term in the language, a core use, such as the one specified above for 'red,' will explain all uses of the term: "All uses of *w* stem from its possession of acceptance property *A(x)*, where *A(x)* gives the circumstances in which certain sentences containing *w* are accepted" (ibid., 45). Thus the task for a philosopher attempting to give a use-theoretical semantics for moral terms is to explain the core use of terms such as 'ought' and 'good.' Once this has been achieved, other terms whose meaning is partially descriptive, and partially moral[20], such as 'murder,' will

[20] In drawing on the common distinction between descriptive and moral terms, I do not intend to beg the question against any metaethical theories. For the naturalistic moral realist, for example, all moral terms will be descriptive terms: Such terms serve to pick out robust properties. (I discuss the contrast between naturalistic

have their meanings explained (in part) on the basis of the meaning of these most basic moral terms. It is important to have a theory that can account for the fact that use of the word 'murder,' as opposed to the purely descriptive term 'killing,' implies that the act in question was wrong, an act that ought not to have been perpetrated. To do so, one must explain the role 'oughts' play in guiding action.

The Core Use of Moral Terms

As noted above, a number of prominent philosophers, including Sellars, Harman, Brandom, and Wedgwood have developed alternative versions of a use-theoretical semantics for normative and moral terms. The essence of all of these views is present in the earliest version of such a theory, the theory of the meaning of 'ought' offered by by Wilfrid Sellars in "Some Reflection on Language Games." Sellars's view is that the meaning of moral terms such as 'ought' is to be explained by the fact that persons who accept sentences containing these terms will act in a certain way. Specifically, if I accept the sentence 'I ought to do X,' then I will try to do X. If I accept the

---

moral realism and other views, and detail the notion of 'robust properties,' in Chapter 4). On such an account, we could draw a distinction parallel to the normative/descriptive distinction based on the difference between those terms that are descriptive and normative (such as 'ought') and terms that are descriptive but not normative (such as 'killing' and 'electron').

sentence 'I ought to finish this paper by tonight,' then I will make an effort to finish this paper tonight. As Sellars writes:

> The motivating role of 'ought' in the first person is essential to the 'meaning' of 'ought.' That is to say, it could not be true of a word that 'it means ought' unless this word had motivating force in the language in which it belongs. (Sellars 1954, 350).

It is the fact that the concept conveyed by the term 'ought' plays such a role in motivation and action that this term has the meaning it has. Normative terms are implicitly defined by the role that acceptance of sentences containing such terms plays in leading one to pursue certain courses of action and avoid others.

That normative and moral concepts play a role in the motivation of action is a central aspect of the account of such concepts offered by Allan Gibbard in his book *Thinking How to Live*. Gibbard has a nice term for such concepts—normative concepts are, in Gibbard's phrase, plan-laden concepts. Accepting a claim that contains plan-laden concepts

> consists in planning to live in a certain way…Some concepts are plan-laden; they are non-descriptive

> concepts that figure specially in thinking how to live (Gibbard 2003, 114).

While Gibbard does not explicitly advocate a use theory of meaning in *Thinking How to Live*, his suggestions cohere well with those that have been made by the use theorists cited above. Gibbard and the use theorists all stress the importance of the relationship between accepting moral claims and the motivation of action.

Perhaps the most well-developed use-theoretical semantics for normative and moral terms is the one that has been presented by Ralph Wedgwood. The account presented by Wedgwood in "Conceptual Role Semantics for Moral Terms" spells out a rule of acceptance for a basic normative term in a more precise fashion. Noting that what will be the best course of action to pursue may depend on the circumstances at a time—for instance, when I should eat a substantial meal will depend on the time of day—Wedgwood offers the following acceptance rule for the term 'is a better thing to do':

> Acceptance of 'B(x, y, me, t)' commits one to having a preference for doing x over doing y at time t.

In this schema, 'B' is a term picking out the concept of being a better thing to do, 'x' and 'y' refer to two

distinct courses of action, 't' is a specific time, and 'me' refers, of course, to oneself.

It is important to note here that Wedgwood, unlike Horwich, thinks that meaning is an intrinsically normative notion, and that one ought to spell out the meanings of words in terms of normative notions such as undertaking a commitment. Disagreeing with other use theorists such as Horwich and Peacocke, Wedgwood writes "I strongly suspect that the notion of 'mastering a rule' can only be explained in partly normative terms" (Wedgwood 2001, 9). As noted above, I do not think that this is correct. This difference between Wedgwood and other use theorists does not matter to the present issue—It is possible to modify Wedgwood's account in a way that does not require that meaning be an intrinsically normative notion. Taking this into account, we can state the acceptance rule for the concept BEING A BETTER THING TO DO[21] as follows:

> Speakers who accept 'B(x, y, me, t)' have a tendency toward having a preference for doing x over doing y at time t.

Once this preference has been formed for doing x over doing y, it will typically cause the individual to act, to do x.

[21] I will use capital letters to refer to concepts in this dissertation.

As noted above, the meaning of this term is given implicitly by the role of acceptance of sentences containing the term in motivation and action.

There is a great deal of detail that could be added here in order to flesh out the account of this basic normative term 'is a better thing to do,' and how to explain other normative terms on the basis of it. This work is outside of the scope of this dissertation. However, the philosophers cited above, particularly Wedgwood and Gibbard, make serious attempts to explain the meanings of a variety of normative terms by starting with accounts of the role of basic, nonmoral, normative terms. It is my goal in this chapter to present an account of the meaning of 'ought' without spelling out in detail all of the particular accounts of moral 'ought's on offer. The hope is that this basic account of the meaning of 'ought' could be appealed to on a variety of metaethical theories. There is still a great deal that needs to be said[22] in order to explain moral 'ought' in terms of this basic 'ought.' To cite one prominent example of how other philosophers have attempted to explain moral 'ought's, on Gibbard's theory, moral 'ought's are explained in terms of the

[22] I am grateful to Catherine Wilson for pressing this point at the defense of this dissertation.

feelings we ought to feel in certain situations. On Gibbard's theory, the moral term 'morally wrong' is defined as follows: an act is morally wrong iff a person ought to feel guilt for doing that act, and others ought to resent that person for doing that act. As stated in this chapter, the basic theory of 'ought' does not require acceptance of any particular account of the meaning of moral 'ought's, including Gibbard's.

An important issue raised by this theory of the meaning of moral terms is the relationship between the semantics of moral terms and moral psychology. To characterize this relationship, I will use the notion of a "belief box," a useful fiction invented by Stephen Schiffer in his characterization of Jerry Fodor's language of thought hypothesis. When a person has a belief that snow is white, for example, we can characterize that belief by claiming that such a person has the mental sentence 'snow is white' in the belief box. We can also use the notion of a "desire box" to characterize a person's desires. When I desire to drink juice, I have the sentence 'I drink juice' in the desire box.

On the account given above of the meaning of 'ought,' whenever a person has the mental sentence 'I ought to do x' in the belief box, then that person will have the sentence

'I do x' in the desire box. Persons who have 'I do x' in the desire box will tend to do x. Thus believing that I ought to prepare better for my classes will result in my desiring that I prepare better for my classes, which will result in my preparing better for classes. The crucial role played by the concept OUGHT in this case and others is the role this concept plays in motivation and action.

What are the Alternatives to a Use Theory?

The predominant alternative to a use theory of meaning is to explain meaning in terms of the theory of reference. There are referential relations between the world and words, such relations require (according to these philosophers) a substantial theory, and we are to account for meaning in terms of these relations. As I will argue below, there are good reasons to be skeptical of the prospects of an account of moral semantics in terms of one of referentialist theories. If a proponent of such a view finds these arguments convincing, it is an open option to adopt the use-theoretic account of meaning for a restricted realm of discourse, such as moral terms. Proponents of referentialist theories have appealed to use theories when their theories have failed to explain the meaning of a certain range of terms. For instance, Fodor, a proponent of a causal/informational referentialist theory of meaning,

appeals to a conceptual role theory in order to explain the meaning of logical connectives such as 'and,' 'or,' and 'not' (Fodor and Lepore 1991). While having two theories of meaning—one for terms generally, another for terms that the theory cannot explain such as logical connectives and moral terms—may be a way to evade the following arguments, it is preferable theoretically to have a uniform account of meaning. For these reasons, the following arguments should be considered part of the overall case for holding a use theory of meaning.

Reference, Meaning, and the Open Question

The trouble with accounting for the meaning of moral terms on a referentialist account is a problem noted by G.E. Moore in *Principia Ethica*. As Moore pointed out, one can provide an expression that is coreferential with 'good,' but it is possible for such a coreferential expression to fail to convey the meaning of 'good.' Say, for the sake of argument, that utilitarianism is correct, and 'good' and 'maximizes happiness' are coreferential. The trouble is that a fully informed person can still meaningfully ask "I know this maximizes happiness, but is it good?" This point is the conclusion of the well-known open question argument.

It is important to note that Moore's point is not simply an instance of the more general phenomenon noted by Frege in "On Sense and Reference." It is a distinct and deeper problem. In situations where a person lacks certain information, that person can always meaningfully ask regarding two coreferential terms whether they in fact corefer: It is possible to meaningfully ask "I know that is Hesperus, but is it Phosphorus?" or "I know he is Superman, but is he Clark Kent?" The disanology between these cases, generally, and the Moore point is that a fully informed speaker—one who is not ignorant of any facts relevant to this situation—could still meaningfully ask regarding 'good' and a purported coreferential term such as 'maximizes utility' whether such terms are coreferential[23].

Many of Moore's arguments in *Principia Ethica*, in particular his arguments concerning the "naturalistic fallacy," are no longer considered convincing. Nonetheless, the open question point itself, regarding the meaning of 'good,' seems quite clear and correct. As Darwall,

[23] I am grateful to Michael Devitt for raising the question of whether Moore's argument is simply an instance of Frege's puzzle. It is important to note here that I do not think that Moore's point, by itself, presents a convincing argument against referentialism. As I will explain in the following section, the problem with referentialism is its failure to account adequately for moral motivation, and a reformulation of Moore's point by Darwall, Gibbard, and Railton illustrates this problem.

Gibbard, and Railton (1992) point out, there is a way of explaining why Moore's point is convincing that does not require an appeal to Moore's own flawed arguments, arguments that are based on "a now defunct intuitionistic Platonism" (Darwall, Gibbard, and Railton 1992, 3). On the explanation of the success of Moore's open question point offered by Darwall, Gibbard, and Railton, the reason we find this point convincing is that it identifies an important fact regarding the motivational role of the concept GOOD. Substituting any coreferential expression for 'good' will fail to explain that goodness is something to be sought. Belief that a certain object is good will motivate a person to act in a certain way toward that object: to bring it about. For this reason, as Gibbard 2003b correctly notes, the Moorean argument[24] shows that the concept of good must be a distinct concept from other concepts that might pick out the same property. As noted above, concepts such as GOOD and OUGHT are plan-laden concepts, and full understanding of these concepts requires showing how propositional attitudes containing these concepts motivate action. The referentialist account fails

[24] I will use the term "Moorean argument" to refer to the Darwall, Gibbard, and Railton adaptation of Moore's argument.

to distinguish plan-laden concepts from coreferential, non-plan-laden concepts[25].

On Sensible Knaves

A possible objection to the account of the meaning of moral terms detailed in this chapter is that it does not answer the objections of philosophers who hold an externalist view of motivation, such as Railton (1986). On an externalist view, one can be fully aware of what one ought to do, what is good, and all other normative facts and still not be motivated to act on the basis of knowledge of such facts. Rocco might be aware that he ought to treat others fairly, but he might nonetheless have no inclination to do so. If Rocco fully understands what he ought to do, but has no inclination to do any of those things he ought to do, he would be, in Hume's term, a "sensible knave." If such "sensible knaves" are possible, then the account of the meaning of moral terms given here is incorrect. One can fully understand terms such as 'ought,' 'best,' and 'good' without having any inclinations toward certain courses of action.

[25] As Michael Devitt pointed out, a referentialist who thinks that there are 'queer' moral facts, in the sense described by Mackie (1977), could hold that belief in moral facts accounts for motivation. Such a referentialist would need to respond to Mackie's influential arguments.

In response to these externalist objections, objections he terms "hyperexternalist," Gibbard (2003b) has convincingly argued that such an account of moral motivation is incoherent. I will briefly summarize the argument Gibbard offers against the externalist. Gibbard asks us to imagine two people, Tweedledum and Tweedledee. Tweedledum has the correct view regarding what he ought to do—for the sake of argument here, imagine that hedonistic utilitarianism is the correct view of what to do. Tweedledee, on the other hand, claims that we ought to be egoistic hedonists. Presumably, most of us would consider Tweedledee, on the basis of these claims, to have a deeply irrational view of morality.

Tweedledee and Tweedledum have long, loud discussions of what is right to do and what is wrong to do, and their arguments never reach a resolution. On the other hand, when we observe what Tweedledee and Tweedledum do, and the choices they make, they end up making identical decisions and taking the same courses of action. In such a case, can we take seriously the externalist claim that, regardless of the identical choices made by Tweedledee and Tweedledum, and their identical actions, that they nonetheless hold radically different views of what they ought to do? It is hard to see why we should take Tweedledee's hedonistic

egoist claims at face value if he ends up acting in the very same way as the utilitarian Tweedledum. It seems that Tweedledee has a certain view of what he ought to do, and such a view is reflected in his decisions and actions. The externalist, on the other hand, would require us to, as Gibbard puts it, attribute "distinctions where they can't be found" (Gibbard 2003b, 12). As Gibbard's argument shows, we cannot plausibly attribute a moral view—a set of beliefs about morality—to a person who does not show any sign of acting in accord with that moral view. It would be especially implausible, as this case shows, to attribute a set of beliefs about morality to a person who is clearly acting in accord with a radically different moral view. The tie between moral motivation and our moral beliefs is closer than the externalist view would claim. If we find regarding Tweedledee both as a convinced hedonistic egoist and as a person who does not act in accord with hedonistic egoism unconvincing, then we ought not to be moved by externalist objections to an account of the meaning of moral terms of the sort described in this chapter.

A referentialist may be unmoved by the intuition behind Gibbard's Tweedledum/Tweedledee case. Such a referentialist might find the logical possibility of coherently attributing a belief system completely at odds

with a person's motivational state and behavior not at all troublesome. Such a referentialist could insist that the only reason philosophers have found such an attribution troubling is that they have taken for a conceptual connection between moral belief and motivation what is instead a contingent correlation. While the referentialist account of meaning cannot explain the relevant facts about motivation and behavior, the referentialist can turn to a theory outside of the theory of meaning—such as an evolutionary theory—to explain the correlation between acceptance of moral claims and moral motivation[26]. I will call this referentialist approach to motivation 'the parallel explanation strategy.'

While the intuition behind Gibbard's Tweedledum/Tweedledee case is widely shared, I will grant for the sake of argument that it does not establish a conceptual tie between the acceptance of moral claims and moral motivation. Even if this were the case, the referentialist account still is at a disadvantage in comparison with the use-theorist who accepts the account of 'ought' given in this chapter. It is important to remember a central methodological point of the use theory of

[26] This objection is due to Michael Devitt. I am grateful to Devitt for pressing this point.

meaning[27]. Horwich makes a crucial distinction, as I noted above, between the core use of a term and other, noncore uses of a term. On the use theory, as Horwich states in *Meaning*, "The overall use of a term stems from its possession of a basic acceptance property" (Horwich 1998a, 44). This basic acceptance property is the core use of the term. It is an important tenet of the use theory to prefer accounts of the core use of a term that will explain the use of a term generally. If we accept this methodological point—as we should, for it is based on an explanatory consideration that is central in all scientific explanation—we ought to prefer accounts of the core use of a term that have greater explanatory power. The use-theoretical account of the meaning of 'ought' has an explanatory advantage over the referential account: it explains facts regarding motivation and behavior that are left unexplained by the referentialist's account of the core use of 'ought.' The account of the core use of 'ought' sketched in this chapter is not only in accord with a wide-spread intuition regarding the tie between moral belief and motivation—this account can better explain the

[27] I am deeply grateful to Horwich for reminding me of this key explanatory point in the use theory and of its relevance to the issue between the use theorist and the referentialist.

many ways we use this term on the basis of a simple account of the term's core use.

CHAPTER 4

MINIMALIST SCHEMAS: TRUTH, PROPERTIES, AND FACTS

Philosophers who have discussed moral matters and their status have not typically restricted this discussion to the matter discussed above, regarding the truth of moral utterances. Such philosophers have also raised ontological questions generally thought to be either closely related to or identical to the questions regarding the truth of moral discourse, such as the question of whether there are genuine moral facts or genuine moral properties. The debates between "moral realists" and "moral anti-realists" are sometimes characterized in terms of these questions. Of particular interest here are the remarks made regarding facts and properties by Simon Blackburn and Allan Gibbard in presenting their expressivist theories, theories that will be discussed in great detail in Chapter 6 of this dissertation.

Blackburn, in his "How to be an Ethical Antirealist," characterizes his "projectivist" view as an "antirealist" one due to the fact that such a theory involves explanations of moral matters that need not make appeal to moral facts and properties:

> This theory is visibly antirealist, for the explanations offered make no irreducible or essential appeal to the existence of moral "properties" or "facts"; they demand no "ontology of morals" (Blackburn 1988, 174).

Gibbard distinguishes, in *Wise Choices, Apt Feelings,* between naturalistic facts, normative facts, and facts of meaning. Gibbard then claims that there are no real normative facts: "In my own picture…[apparent] normative facts, will come out, strictly, as no real facts at all…" (Gibbard 1990, 23). In his recent book *Thinking How to Live*, Gibbard characterizes expressivism as denying that there are any *distinctive* normative states of affairs or normative facts[28]:

---

[28] Gibbard does grant that it is possible to hold there are normative facts on a minimalist conception of facts, a position argued for by Horwich 1993 and defended further in this chapter. "Suppose, instead, that minimalists are right for truth, for facts, and for belief; there is no more to claiming 'It's true that pain is bad' than to claim pain is bad; the fact that pain is bad just consists in pain's being bad; to believe that pain is bad is just to accept that it is. Then it's true that pain is bad and it's a fact that pain is bad—so long as, indeed, pain is bad" (Gibbard 2004, 182-183). Gibbard expresses doubts regarding the possibility of a contrast between such minimalist facts and a more robust species of fact. "Are these just pseudo-facts, incapable of real truth or falsehood?…I still weasel: I say that I need to understand the questions. Explain to me "real facts," "substantial truth," and "genuine belief," and I can think how to

> There is no such thing as a specifically normative state of affairs; all states of affairs are natural…Then, clearly if my quasi-realism is correct, there aren't distinctively normative facts, only naturalistic facts (Gibbard 2003, 181).

This discussion of moral ontology by Blackburn and Gibbard takes place regardless of the fact that facts and properties, at least at first sight, without further explanation, are seemingly rather strange, not terribly well understood sorts of entities. Philosophers have given a number of theories of facts and properties that have failed to resolve these worries regarding their status. It will be necessary to settle these general issues regarding the nature of facts and properties before addressing the specific issues that have been raised regarding moral properties and facts.

In the spirit of the deflationary theories of truth detailed in Chapter 1, we can turn to a deflationary theory of facts and properties in order to clarify this issue. As Paul Horwich has noted in his paper "Gibbard's Theory of Norms," "…[P]arallel accounts [to the minimalist truth schema] will hold of notions such as 'fact' and 'property'

---

answer." (ibid., 182). The aim of this chapter is to provide such an explanation, and to defend it.

that are intimately related to 'truth'" (Horwich 1993, 73). In this chapter, I will spell out in detail and defend the account of properties and facts presented by Horwich in order to give a clear, well-defined account of these notions. This account of facts and properties will provide a framework for the discussion of meta-ethical theories and their metaphysical underpinnings in subsequent chapters.

One might worry that such a deflationary account of facts and properties would not be capable of properly characterizing the sort of disputes mentioned above, disputes regarding the existence or nonexistence of genuine moral facts and genuine moral properties. In the end of this chapter, I will show that, within this framework, the traditional disputes over the existence of moral facts and properties can be characterized in a far clearer fashion than is usually the case in discussion of these matters. That this is the case would be surprising to philosophers who hold the view I will discuss in Chapter 6, the view that there is a tension or incompatibility between an expressivist approach to meta-ethics and deflationism regarding truth.

Facts

As noted above in the **Inflationism and Correspondence** section of Chapter 1, certain traditional versions of the

correspondence theory of truth have defined true propositions (or utterances) as those propositions (or utterances) that correspond to the facts. This characterization is, for reasons discussed above, incomplete. In order to complete the theory, the correspondence theorist would have to tell us what a fact is. Is it some sort of object? In what way does this object resemble, or correspond, to a proposition? Is a fact some sort of (in J.L. Austin's phrase) "linguistic *Doppelgänger*" (Austin 1970, 123)?

Is there any need to posit the existence of these kinds of facts as objects over and above the concrete objects we are already familiar with? As Quine (1960) points out, the sentences 'Fifth Avenue is six miles long' and 'Fifth Avenue is a hundred feet wide' would presumably be made true by different facts on the correspondence theory. However, in order to explain fully what makes such sentences true, we need not posit entities such as facts over and above objects. We can account for the truth of these sentences in terms of a single, more familiar object—Fifth Avenue:

> Our two sentences last quoted are true because of Fifth Avenue, because it is a hundred feet wide and six miles long, because it was planned and made that

way, and because of the way we use our words; only indirection results from positing facts, in the image of sentences, as intermediaries (Quine 1960, 247).

Thus it seems that facts, as traditionally conceived of by the correspondence theorist, are both obscure and theoretically unnecessary.

An argument due to Frege (1918-19), Church (1956), and Davidson (1984) raises a further, serious difficulty for proponents of any version of the correspondence theory spelled out in terms of facts[29]. This argument is often called the "slingshot argument." If, according to the slingshot argument, a statement is made true by corresponding to a particular fact, then all true statements are made true by corresponding to the same fact. Instead of there being individual, distinct facts that make different statements true, there could only be one fact, in Davidson's term, "The Great Fact."

The slingshot argument is based on the following two assumptions: (1) If a sentence *S1* corresponds to a certain fact *F*, then any logically equivalent sentence *S2* that we substitute for *S1* will correspond to the same fact *F*; (2) If a sentence *S1* corresponds to a certain fact *F*, then the

[29] The most comprehensive discussion of the slingshot argument and its implications is Stephen Neale's *Facing Facts* (2001)."

substitution of a singular term *T1* (contained in *S1)* for a coextensive singular term *T2* will yield a sentence *S2* that corresponds to the same fact *F.*

Take as our example of a statement that happens to "correspond to a fact" the true sentence *S1*:

S1: The Yankees won the 1998 World Series.

By assumption, this statement is made true by the fact that the Yankees won the 1998 World Series. Appealing to assumption 1, if we substitute for *S1* the sentence *S2*, which is logically equivalent to *S1*, we have substituted a sentence that corresponds to the same fact F.

S2: (the x such that x is identical with George W. Bush and *S1*) is identical with (the x such that x is identical with George W. Bush)

Given that *S1* is identical in truth value with sentence *S3*,

S3: The Red Sox won the World Series in 1918.

we can substitute *S3* for *S1* in sentence *S2* without changing the reference of the singular term in the subject position. Thus, appealing to assumption 2, we see that *S4* corresponds to the same fact as the fact to which both *S2* and *S1* correspond:

S4: (the x such that x is identical with George W. Bush and *S3*) is identical with (the x such that x is identical with George W. Bush).

Given that *S4* is logically equivalent to *S3*, we can make another appeal to assumption 1 to establish that *S3* corresponds to the same fact as *S4*, as well as the same fact to which *S1* and *S2* correspond.

To reiterate, the argument is as follows:

SS1: S1 corresponds to fact F (assume for the sake of reductio)

SS2: S2 corresponds to fact F (SS1, assumption 1)

SS3: S4 corresponds to fact F (SS2, assumption 2)

SS4: S3 corresponds to fact F (SS3, assumption 1)

Thus, if S1, "The Yankees won the World Series in 1998" and S3, "The Red Sox won the World Series in 1918" are made true by corresponding to a fact F, they are made true by corresponding to the very same fact. Repeated appeals to this line of argument would establish Davidson's conclusion that, if statements are made true by corresponding to facts, then all statements correspond to one and the same "Great Fact."

Davidson takes the moral of this argument to be that we should reject philosophical theories spelled out in terms of facts: "The strategy of facts, against which I have been inveighing, is something else: a philosophical theory, and a bad one" (Davidson 1984, 50). Quine and Davidson are not the only philosophers whose views lead

them to rejection of talk of facts. Other versions of the correspondence theory of truth, such as the version of Tarski's theory proposed by Field (1972), make no appeal to facts in order to explain the relationship between language and the world. As Davidson (1984) points out, if the Tarskian theory of truth is correct, a full definition of truth can be given by appeal to principles regarding the denotation of referring terms and the satisfaction of predicates. No appeal to facts is required on such an account.

Non-correspondence inflationary theories of truth including the pragmatist theory, the coherence theory, and the unanalyzable quality theory also seem to require no appeal to this correspondence notion of facts. For, if true propositions are defined as propositions that are useful to believe, then we need not make any appeal to facts in order to explain the nature of truth. We need only appeal to an account of practice that will tell us which propositions are more or less useful. Also, none of the deflationary theories of truth makes appeal to the notion of a fact in explaining the role of the truth term. It would seem that the only theory of truth that gives facts an essential role to play is the version of the correspondence theory that is susceptible to the slingshot

argument. Should we then follow Quine and Davidson in rejecting facts *tout court*?

Davidson's counsel to regard talk of facts as part of a useless philosophical theory overlooks that talk of facts, in ordinary contexts, serves important purposes. For example, we certainly wouldn't want to follow the strategy of rejecting facts to the absurd consequence of refusing to recognize as meaningful statements such as 'Voters will decide in this election based on the facts on the ground in Iraq,' or 'Successful scientific theories are based on the observed facts.' Davidson seems to recognize this to some degree when he calls "talk of facts…harmless," but he fails to offer any account of facts (Davidson 1984, 50).

A way to properly characterize the way the notion of fact is employed in ordinary contexts without making appeal to any strange, theoretically unnecessary entities is provided by a schema similar to the minimalist truth schema discussed in Chapter 1. Truth is characterized in terms of acceptance of instances of the following schema:

<p> is true if and only if p.

Whenever we accept that p, we also accept that it is true that p. It is also the case that whenever we recognize that a certain proposition, <p>, is true, we are inclined

to claim that it is a fact that p. Given that it is true that electrons have negative charge, it is a fact that electrons have negative charge.

Given that this is the case, we can characterize our notion of fact in terms of acceptance of the following schema:

That p is a fact if and only if *p* (Horwich 1993, 74)

We can appeal to this minimalist fact schema in order to see that the notion of fact plays a generalizing role similar to the generalizing role played by the notion of truth. Talk of the facts in a particular domain is a shorthand allowing us to refer to a number of true propositions without stating them one at a time. For instance, take the sentence "Voters will decide in this election based on the facts on the ground in Iraq." What this means is that, for all propositions <p>, if <p> is a proposition regarding the current situation in Iraq, and it is a fact that p, then the voters will make decisions in the election based on the belief that p. By the equivalence schema, this amounts to the following: for all propositions <p>, if <p> is a proposition regarding the current situation in Iraq, and p, then the voters will make decisions in the election based on the belief that p.

This alternative, minimalist account of facts allows us to properly characterize the notion of facts used in ordinary contexts as well as to diagnose the error that led to the conclusion of the slingshot argument. The presupposition of the correspondence theory that truth involves correspondence to facts is rejected by deflationists generally, including the minimalist. A minimalist account of truth, by providing an account of truth that makes no appeal to correspondence in its explanation of truth, is not damaged by the slingshot argument.

The trouble the slingshot presents for correspondence theories is that one cannot take the notion of fact to be an *explanatory* notion, one that plays an important role in a theory of truth, if the slingshot argument is successful. For how could correspondence to facts explain why certain statements are true, if it is the case that when any statement corresponds to a particular fact, it corresponds to all other facts as well? The minimalist account of truth and facts, in stark contrast, does not account for the truth of propositions in terms of a correspondence relation between propositions and facts. Thus the slingshot considerations regarding the relation of correspondence between statements and facts does not

present any problem whatsoever for the minimalist account of truth and facts.

If we accept a minimalist account of facts, would it be possible for us to deny the existence of moral facts? Given the arguments presented in Chapter 2 for attributing truth to moral statements, there seems to be no comfortable room for such a position. For, given that there are good reasons to attribute truth to moral utterances, and given that saying that a certain proposition that p is true is trivially equivalent to saying that it is a fact that p, we have good reason to claim that there are moral facts. Blackburn, when denying the existence of moral facts, must have some more "robust" notion of fact in mind when denying that projectivism makes appeal to such facts. That this is so is clear from Blackburn's own claims that his projectivist theory is fully consistent with attributions of truth to moral utterances (Blackburn manuscript).

What makes a fact a "robust" fact can be a bit obscure, and one often finds philosophers explaining the "robustness" of facts in a circular fashion, or appealing to words with similar meaning such as "real" or "substantive" to attempt to make such a distinction. We could plausibly, in light of his discussion of the metaphysics of morality in other contexts, take Blackburn

to be denying that moral facts are robust insofar as they are not appealed to in causal explanations. This is a way one could characterize a fact as "robust." I will discuss this distinction between robust and nonrobust facts in much more detail later in the chapter. In order to provide an account of robust facts that will show how different positions on moral ontology can correctly be characterized, I will turn to Blackburn's other formulation of his view, as a view that does not require making any appeal to moral properties.

Properties

It does not seem possible, *prima facie*, to explain the difference among various positions on moral ontology in terms of the account of facts presented in the previous section. Do we have a clear understanding of properties that will allow for such a distinction? What would it mean to either hold or deny that there are moral properties such as goodness, justice, and virtue?

There is no single, universally accepted view among philosophers about the nature of properties. In fact, disputes over the nature of properties have a great historical lineage, dating back at least to Parmenides, Plato, and Aristotle. In the next section, I will sketch three major philosophical theories of the nature of

properties: platonism, nominalism, and the trope theory. With these theories as the background, I will explain why it is especially difficult to explain, as Blackburn attempts to, a position on the ontology of morals in terms of properties. In the following section, I will propose a deflationary alternative to the three major theories of properties, spelled out in terms of a minimalist property schema analogous to the minimalist truth schema. In the final section of this chapter, I will appeal to a distinction between robust and nonrobust properties that will allow the deflationist about properties, facts, and truth to account for the metaphysical disputes at issue within meta-ethics and other fields that raise similar ontological questions.

Theories of the Nature of Properties

The central issue theories of properties are intended to address is the issue in virtue of what is a certain object with a property F an object with such a property. In virtue of what are all chairs objects that have the property of being chairs, and in virtue of what are all persons objects that have the property of being persons? When we have two dogs, Fido and Spot, and both dogs have the property of being dogs, how is it that these two different animals have one and the same property? This

problem is the traditional "one over many" problem posed by Plato.

To understand fully the various commitments of the philosophical theories of properties, it is useful to appeal to a characterization of these views due to Keith Campbell. Campbell makes a distinction between entities that are universal and those that are particular, and he makes a further distinction between entities that are abstract and those that are concrete. Keeping these two distinctions separate is quite important. Entities are universals if they are "unrestricted in the plurality of different locations in space-time at which they may be wholly present" (Campbell, 477). A particular, unlike a universal, can only be at a certain place at a certain time. Abstract entities are those "got before the mind by abstraction, that is, by concentrating attention on some, but not all, of what is presented" (ibid., 478). No such process of abstraction is required in order to consider a concrete object, like a table, chair, book, or computer. To think of the whiteness of the computer, however, or the hardness of the table, one must concentrate on these particular aspects.

The platonist theory postulates the existence of abstract universal objects, a property (in the case of Fido

and Spot) of being a dog, which all objects which have this property exemplify. There is a single entity that is the property F, and this universal F is exemplified by all objects that are F. In addition to individual concrete chairs, there is an individual abstract object that is the property of being a chair. It is the relationship between the abstract object of being a chair and the individual chairs that is supposed to explain why it is that all chairs are chairs.

The trope theory differs from platonism insofar as it is, unlike nominalism, committed to the existence of properties but not to the existence of abstract universals. Rather, the trope theorist holds that properties are abstract particulars. There is not a single abstract object, the property of being a chair, that accounts for why all chairs are chairs. Instead, each chair is supposed to contain an instance of a trope of chairhood, which is some part of the chair. All of the other aspects of the chair—its hardness, its being made of wood, its shape—are also tropes, and the chair itself is a collection of these tropes.

The nominalist theory balks at the postulation of either of these odd sorts of abstract entities, abstract universals or abstract particulars, and claims that there

are no such things as properties. Instead, there are only predicates and objects. The proper approach is to eschew talk of an object having the property of being F, and just to note that the object is F. There is nothing to say over and above the predication of 'F' of the object. As Michael Devitt has noted, the nominalist offers an answer to the question of the one over many problem, "In virtue of what are *a* and *b* both *F?*" The answer given by the nominalist is simply that this is true in virtue of the following:

(1) *a* is *F;*

(2) *b* is *F* (Devitt 1980, 95).

As Devitt notes, this cannot be all there is to be said regarding properties on a adequate nominalist account, for there are a number of meaningful expressions purportedly referring to properties that cannot be explained by claims such as (1) and (2) alone, such as "red is a color."

What Would Commitment to Moral Properties Be?

How are we supposed to read any claim denying the existence of moral properties? The trouble is that it seems that one needs to appeal to one or another of these philosophical theories of properties in order to cash out this denial of existence, and doing so will require becoming entangled in this debate over the nature of properties.

Take, for example, the nominalist theory. The nominalist denies that there are properties. Instead, there are only predicates and objects to which predicates apply. If this is correct, then as Michael Devitt (2002) points out, the nominalist is automatically committed to denying the existence of moral properties. Denial of the existence of moral properties is just a particular instance of the nominalist's general position on the nonexistence of properties.

On the other hand, if we are to understand positions that claim moral properties do indeed exist, then we would have to appeal to either the platonist theory or the trope theory in order to fully spell out what this commitment amounts to. It is far from clear, however, that all philosophers who are willing to claim that properties such as goodness, justice, and virtue genuinely do exist would also be comfortable claiming that there are either abstract universals or abstract particulars. One could coherently imagine a philosopher being perfectly comfortable claiming virtue and justice exist without feeling particularly attracted to either platonism or the trope theory. For these reasons, as noted above, assertion or denial of the existence of moral properties would be tangled into these vexed issues.

A further difficulty with any of these traditional theories of properties is that none of them provides us with the resources to hold that certain predicates do in fact pick out properties whereas other predicates do not. If we are to have a dispute over whether or not there are in fact moral properties, these theories do not present us with the resources to make such a distinction[30].

A Deflationary Account of Properties

It is important to notice that all three of the traditional accounts of properties discussed above are deeply unsatisfying, and result in either regarding the existence of properties as a mysterious matter or revising our ordinary practice of attributing properties[31]. The platonist makes it difficult to understand how it is even possible that objects could have properties. We are driven to a philosophical paradox by the account given of the

[30] This is not to suggest that all platonists, nominalists, and trope theorists accept the simple versions of the (respective) theories sketched above. Any property theorist could supplement her theory with an account that distinguishes the predicates that pick out genuine properties from those that do not. Armstrong, for example, could make such a distinction based on his theory that the properties that are genuine are those that figure into scientific laws. I am grateful to Michael Devitt for pointing this out.

[31] The criticism of the traditional account of properties in this section owes a great deal to Wittgenstein's account of philosophical problems in the *Philosophical Investigations*, and to Paul Horwich's discussion of Wittgenstein's account in courses given at the CUNY Graduate Center.

relationship between objects and properties. If have a property, and being in a relation, are explained by the postulation of properties and relations, then we will need to account for the relation that binds the property to the object. The trouble that presents itself is that this relation of exemplification, the relation that binds property to object, itself stands in need of explanation. Postulation of a further relation to relate property, object, and exemplification will result in infinite regress, a philosophical paradox.

The trope theorist, on the other hand, presents a view that is deeply revisionary of our ordinary view both of properties and objects. We would normally say that there is only a single property of being a dog, or a single property of being a table. On this view, however, there are many properties of being a dog, as many properties of being a dog as there are dogs. Where there is a dog, there is a trope of being a dog. Perhaps even more difficult to understand is the position of trope theorist on the existence of objects. For Campbell, the basic kind of entity is a trope, and objects are collections of such tropes.

> On the view that tropes are the basic particulars, concrete particulars, the whole man and the whole

> piece of cloth, count as dependent realities. They are collections of co-located tropes, depending on these tropes as a fleet does upon its component ships (Campbell, 479).

Thus the trope theorist makes even ordinary objects such as tables and chairs seem obscure and mysterious. What exactly are the tropes that this chair consists of? Where does one trope end and the other begin? How do all of these tropes "hang together"?

The nominalist, in reaction to the mysteries presented by the views of the trope theorist and the platonist, denies the existence of properties altogether. This denial, unless supplemented with an account of ordinary talk, fails to make sense of our ordinary discussion of properties. Devitt (1980), citing Arthur Pap and Frank Jackson, correctly notes that:

> We all assent to, express, believe statements like the following:
>
> > …Red resembles orange more than it resembles blue;
> >
> > …Red is a colour;
> >
> > …He has the same virtues as his father;
> >
> > …The dresses were of the same colour.

> …[T]hese statements seem to require the existence of properties for them to be true (Devitt 1980, 99).

Without, as Devitt notes, a method for paraphrasing ordinary talk of properties, the truth of these statements is left unexplained.

The error that is at the source of all of these various theories of properties is an error of the sort that Wittgenstein noted in the *Philosophical Investigations* as characteristic of philosophical perplexities. Each of these positions results from thinking of properties as being analogous to concrete objects, and expecting talk of properties to function in a way similar to discussion of concrete objects. This is a result of the use of noun phrases such as "the property of being blue" or "the color of this computer" in discussing properties. When we talk of the property P as being a thing that in virtue of which certain objects have the property P, or share the property P, we are misled into thinking that there is some explanatory relationship between something similar to a concrete object, the property P, and the things that have this property. The platonist and trope theorist are led by this analogy to postulate the existence of quite strange objects, objects of the sort one might reasonably doubt exist. The nominalist goes quite far in the other

direction, denying the existence of properties and offering revisions of our ordinary talk of properties.

Is there an account of the notion of property that accords better with our ordinary linguistic practice? Is there some noncontroversial account of properties to which we could turn in order to characterize the kinds of disagreements over ontology that have motivated various meta-ethical positions?[32] If we were to attempt to characterize the use of the term 'property' in ordinary, nonphilosophical contexts, we would find that there is little difference between attributing a property F to an object x and saying that x is F. Is there a circumstance in which we would say that water has the property of being wet, but deny that water is wet? Is there a circumstance in which we would say that George W. Bush has the property of being a Republican, but deny that George W. Bush is a Republican? These entailments work in the opposite direction as well: There is little room for saying that Susan is a redhead but denying that Susan has the property of being a redhead.

---

[32] I am grateful to Jonathan Cohen for his suggestion, during an E-mail discussion over my paper "Linguistics, Psychology, and the Ontology of Language," that a noncontroversial account of properties might serve the purpose of properly characterizing ontological disputes. The account of properties I am appealing to here is, as noted in the text, due to Horwich (1993).

Noting this uncontroversial fact regarding the notion of property, we can characterize our use of the term 'property' by the following schema, suggested by Horwich:

> For any object *x*, *x* has the property of *being F* if and only if *x* is *F*. (Horwich 1993, 74)

There is little room, if properties are understood in this sense, for disagreement over whether a certain property is being attributed to an object by a speaker. For, whenever a speaker predicates F of an object, it just trivially follows that this speaker is attributing the property of being F to that object[33].

Seeing that this is the notion of property employed in ordinary contexts, we once again see where the three traditional accounts of properties discussed above went wrong. The platonist and trope theorist we led to posit entities due to the feeling that there was a deep mystery regarding how a certain object could have a property F.

---

[33] As noted above, the account of properties presented here is due to Horwich (1993). A similar account of properties is discussed by Stephen Schiffer in his book *The Things We Mean*. According to Schiffer, "It is a *conceptual truth* that if Lassie is a dog, then Lassie has the property of being a dog" (Schiffer 2003, 61). I find Schiffer's characterization of properties as "pleonastic entities" secured by "something-from-nothing" transitions a bit misleading. The point is not that properties are created or brought into existence by appeal to the minimalist property schema. Rather, as with truth, the point is that it is a mistake to say anything about properties over and above what is stated by this schema.

The nominalist denies the existence of properties for the same reason: Given that properties are so mysterious, and the only way to account for this mystery is to postulate the existence of a strange sort of object, properties must not exist at all.

In order to avoid confusion, it should be noted that the minimalist account of properties is not a nominalist view—it is a view that, like platonism, introduces properties into our ontology. Unlike platonism, the minimalist account claims that there is nothing further to be said about properties beyond what is spelled out by the trivial property schema. It is through the introduction of properties into our ontology that the minimalist account is able to avoid the problems that plagued the nominalist regarding statements quantifying over properties (such as "red is a color.")

As was the case with facts, as noted above, while the deflationary account of properties presented here resolves problems that have plagued previous accounts, it seems at first sight ill-suited to characterize Blackburn's position and the positions of philosophers with similar views. For Blackburn is willing to endorse unqualified moral claims. For example, in his "How to be an Ethical Anti-Realist," the essay quoted earlier in this chapter, he defends the

claim that his "projectivist" view is capable of explaining how one can make the assertion that cruelty is wrong. This assertion, along with the property schema, would seem to issue in the result that a "projectivist" ought to endorse the claim that cruelty has the property of being wrong as well. There seems to be no room on this view for positions that endorse moral claims but deny the existence of moral facts and properties.

Robust and Nonrobust Properties

In light of these considerations, the best approach for an expressivist to take is not to deny the existence of moral facts and properties. Such denials, insofar as they were to be taken at face value[34], are not well motivated in light of the deflationary accounts of facts and properties mentioned above. If these denials are thought of as essential to an expressivist account, then expressivism stands in need of revision, as Horwich (1993) suggests. If, on the other hand, such denials are clearly not an essential element of expressivism, as I will argue in Chapter 6 of this dissertation, then certain statements by expressivists regarding facts and properties were ill-

[34] It is important to note here that Blackburn's denial of the existence of moral facts and properties, cited above, places the terms 'fact' and 'property' in scare quotes. Perhaps this reflects Blackburn's own suspicion of the traditional discussion of these two notions.

motivated. In either case, it is clear that the expressivist ought not to deny the existence of moral facts and properties. The root of such denials was holding one or another of the theories of properties discussed above, and the problems with such accounts should be clear by now.

There is, however, a difference on the issue of ontology between the expressivist and her meta-ethical antagonists. Such a distinct position on commitment to moral ontology is central to the case that is made in favor of expressivism. Expressivism, traditionally, is supposed to be an appealing view due to its lack of commitment to a strange moral ontology. In order to properly characterize the differences between various meta-ethical positions on ontology, it is necessary to add to the deflationary account of properties presented above a way to distinguish robust properties from nonrobust properties. The expressivist, unlike "moral realists," does not think that there are robust moral properties[35].

[35] David Copp (2001) also appeals to a distinction between robust and nonrobust properties in order to distinguish expressivism from moral realism. He leaves the specification of the distinction between robust and nonrobust properties quite vague: "In the first place, then, moral realism holds that the chief semantic role of the moral predicates is to refer to moral properties, such as rightness, wrongness, virtuousness, viciousness, and so on. Second, it holds that these properties have the same basic metaphysical status as ordinary nonmoral properties,

What is the mark of a "robust" property? What are expressivists and projectivists claiming when they make no appeals to "robust" moral properties? To find an adequate mark of the distinction between robust and nonrobust properties, it is important to consider the views that stand in contrast to expressivism, namely the so-called "moral realist" views. There are two main strands of moral realism: naturalistic moral realism and non-naturalistic moral realism.

The naturalistic moral realism of Richard Boyd and Peter Railton identifies moral properties with properties that are appealed to in scientific explanations. One mark of the properties that figure into scientific explanations is that such properties play a causal role. Just as properties such as being an electron or being a pain state figure into physical or psychological explanations, properties such as being right and being good figure into moral explanations on the naturalistic realist account. Thus, on the naturalistic moral realist account, moral properties are identified with properties that play a causal role. The minimalist theory of properties spelled

---

whatever that is" (Copp 2001, 4). A central aim of this chapter is to spell out the relevant "whatever that is" in a way that distinguishes the expressivist view of moral properties from the view held by moral realists regarding the nature of such properties.

out above provides a simple and uncontroversial account of when this is the case: when we use moral predicates such as 'right' and 'good' in correct causal explanations, we are appealing to causally efficacious moral properties.

The difference between the naturalistic moral realist and the non-naturalistic moral realist is that non-naturalists will deny that moral properties can be identified with any of the properties that are invoked in scientific explanations. Nonetheless, non-naturalists such as Moore (1903) and Shafer-Landau (2003) do not claim that moral properties exist in some mysterious realm outside of the natural world: Both Moore and Shafer-Landau claim that moral properties supervene on naturalistic, scientifically respectable properties. This, however, is an uncontroversial view, one that any philosopher who holds that moral properties exist should hold.

One major mark of the non-naturalist view, one that makes a further distinction between naturalistic moral realism and non-naturalistic moral realism, is the account of moral epistemology offered by the non-naturalist. Moore and Shafer-Landau believe that moral knowledge is a priori knowledge. What accounts for the existence of a priori moral knowledge?

The non-naturalist could account for the existence of a priori moral knowledge in terms of a causal relation between moral properties and beliefs regarding such properties. In fact, this is how Moore and Shafer-Landau proceed. In Moore's case, he appeals to an intuitionist account: We are aware of moral properties and facts through the process of intuition. While we cannot provide any analysis of 'good' in terms of some other, naturalistic property, we have knowledge through intuition of the simple, unanalyzable property of goodness.

One Shafer-Landau's account, knowledge of moral matters is explained on a reliabilist model. What makes us justified in believing moral claims is the fact that our process of forming moral beliefs is a reliable process, one that correctly tracks moral facts. What would account for the reliability of this process of moral belief formation? This reliable process can be explained in terms of the relationship between such moral beliefs and moral properties.

Thus on either of these views of moral epistemology, Moore's view as well as Shafer-Landau's, we need to appeal to a causal relationship between moral properties and moral beliefs. Thus the strand that runs through the moral realism of the non-naturalist and the naturalist is the

appeal to causally efficacious moral properties. Thus we can characterize the view denied by expressivists, the view that there are robust moral properties, as follows:

> **Robust Properties**: A property is robust iff such a property figures into causal explanations.

We can further define a notion of robust fact as follows:

> **Robust Facts**: A fact F is a robust fact iff such a fact involves the attribution of a robust property.

Thus the central metaphysical component of expressivism is best characterized as the view that there are moral facts and moral properties, but there are no robust moral properties, and no robust moral facts[36]. As Dreier (2004) points out, there is a central explanatory difference between views such as Blackburn's and Gibbard's and the views of realists such as Boyd, Railton, Moore, and Shafer-Landau: the realist invokes a robust moral property in causal explanations, whereas the expressivist does not. The bearing of the issue of whether there are robust moral properties on the issues of moral truth and moral psychology will be considered further in the next chapter.

---

[36] In light of this account of facts and properties, it should be clear that Ronald Dworkin's (1996) claim that there is no way to draw a distinction between moral claims (such as 'Allowing people to die in poverty is wrong) and claims about the status of moral facts (such as 'There are no robust moral facts') is incorrect.

CHAPTER 5

MORALITY, METAPHYSICS, SEMANTICS, AND PSYCHOLOGY

Consider the following three questions:

4. Is it possible for utterances in moral discourse, such as 'Rape is wrong' and 'Great inequalities in the distribution of wealth are unjust,' to be true or false?
5. Are there moral properties, such as goodness and justice? Are such properties ones that figure into scientific or causal explanations?
6. Are moral utterances such as 'Rape is wrong' expressions of moral beliefs, or merely of non-doxastic mental states such as desires or emotions?

It is striking that, while these seem intuitively to be distinct questions, which may be answered yes or no in a variety of combinations, the contemporary literature in metaethics often assumes that an answer of "yes" to one of these questions will force one to answer "yes" to all of them, and an answer of "no" to one of these questions will similarly force one to answer "no" to all of them. As I will discuss in the section below, **The Cognitivism/ Noncognitivism Distinction**, the received view in philosophy

seemingly is that the only logically possible combinations of answers to these questions are, one the one hand, "Yes, yes, and yes," and on the other hand, "No, no, and no." In light of the discussion in the previous four chapters, I will show that these three questions can be answered in a myriad number of ways, hence it is a vast oversimplification to assume that there are only two possible combinations of answers to these questions.

In the following three sections, I will discuss how to answer these questions in turn. In the first section, I will repeat the possible positions on moral truth, as described in Chapter 2. In the second section, I will reiterate the distinction made in Chapter 4 between robust and nonrobust properties, and show that one's position on commitment or lack of commitment to robust moral properties is not automatically determined by one's answer to the question of moral truth. In the third section, I will discuss moral psychology, and the issue of whether moral commitments consist of beliefs, desires, or some combination of the two. Then I will argue that these issues of moral psychology are distinct from the issues of moral truth and commitment to robust moral properties. In the final sections of this chapter, I will relate these distinctions to the discussion of "cognitivism" and

"noncognitivism" in the philosophical literature. I will then describe the variety of possible meta-ethical positions, including a number of logically possible positions that are typically not recognized as such.

It is important to note here that I am not suggesting that questions 1-3 are the only significant or relevant questions in meta-ethics. These are, however, the questions that have the greatest bearing on the issues of whether deflationary theories of truth are consistent with expressivism and constructivism (to be discussed in Chapters 6 and 7). The following issues are outside of the scope of this dissertation: how well various meta-ethical theories account for moral justification; whether certain theories can or cannot account for the possibility of moral knowledge; and to what extend such theories can fully account for our sense that moral justifications are objective. To reiterate, while I think these questions are important and significant, they are questions regarding how successful various meta-ethical theories are, not whether such theories are consistent or inconsistent with particular theories of truth. I am focusing on the latter question.

Moral Truth

The central question in meta-ethics concerning the truth of moral utterances is the one that was posed in Chapter 2 of this dissertation: are moral utterances straightforwardly true or false? There are two possible answers to this question—either moral utterances are capable of being straightforwardly true or false or they are not. To take the former route is to hold what I have called the Attribution Thesis. To take the latter route, one would have to hold either the Denial Thesis, according to which truth or falsehood cannot be attributed to moral utterances, or the Anomaly Thesis, according to which only non-straightforward truth or falsehood can be attributed to moral utterances. As I stressed in Chapter 2, the most plausible route for any theorist is to hold the Affirmation Thesis.

Moral Metaphysics

The question of the metaphysical status of morality is the ontological question posed in Chapter 4 of this dissertation: are there robust moral facts and robust moral properties? There are two possible positions here—either there are robust moral properties and facts, or there are not robust moral properties and facts. As I noted in Chapter 4, the question of whether or not there are robust

properties and facts in morality is best accounted for in terms of causation: if one claims that there are moral properties that play a causal role, such a view is one according to which there are robust moral facts and moral properties. Expressivist accounts in meta-ethics make no such appeal to causally efficacious moral properties, thus such philosophers need not hold that there are robust moral properties.

As I spelled out these positions in Chapter 4, it should be clear that one's stance regarding robust moral properties and facts is not automatically settled by answering the question of whether truth ought to be attributed to moral utterances. On any of the deflationary conceptions of truth detailed in Chapter 1, it is clear that truth could be attributed to an utterance such as 'Killing is wrong' regardless of whether wrongness is a robust property, i.e. a property that figures into causal explanations. What does determine whether or not truth should be attributed to utterances of 'Killing is wrong' is whether killing is wrong. If wrongness is predicated of killing, i.e. when (according to the property schema explained in Chapter 4) the property of wrongness is attributed to killing, it is not necessarily the case that a robust property is being attributed to killing. It is a

separate matter whether we can appeal to the wrongness of killing to explain whether the wrongness of such acts is what causes people not to commit them, or whether it is the wrongness of killing that causes us to believe that killing is wrong. One could hold that killing is wrong yet deny any such causal claims. So, if the property of wrongness is attributed to killing, then (by the deflationary schema), there is a true attribution of wrongness of killing. This does not, for reasons discussed above, imply that wrongness need be a robust property.

Moral Psychology

The major question regarding moral psychology is the question of what sort of mental events are involved in moral commitment and moral decision-making. Can one, as Hume (2000) famously claimed, never be motivated by a belief? If so, are all moral commitments and moral decision essentially emotions or desires? Or rather, as Kant (1996) held, are commitments and decisions that arise from desires non-moral? Are all moral commitments and moral decisions thus essentially beliefs? Or, is it possible that moral commitments and moral decisions are some combination of both beliefs and desires?

Is this issue of moral psychology, the issue of the nature of moral commitments, an issue that is inseparable

from the issue of moral metaphysics? Is it possible to combine distinct views on moral psychology with various views on metaphysics? Can we only have moral beliefs if moral utterances are straightforwardly true and there are robust moral properties? Would claiming that there are merely moral desires but no moral beliefs preclude the attribution of straightforward truth to moral utterances, and lead us to deny the existence of robust moral properties?

There are four possible views one could hold regarding the role of beliefs and desires in moral decision making: (1) Moral decisions consist in desires alone (2) Moral decisions consist in beliefs and desires (3) Moral decisions consist in beliefs. I will consider views (2) and (3) first. On either of these views, it is clear that holding such a view of moral psychology does not automatically settle the issues of moral truth or moral metaphysics. It is possible to have beliefs about both robust properties and nonrobust properties. Consider other properties that may or may not play a causal role, such as color. It is a subject of debate in philosophy whether colors are objective properties, properties that can be invoked in explanations of color perceptions, or whether color properties are subjective properties, projections

from perceivers onto the world. In either case, it is clear that we can have beliefs about colors: We can believe a certain object is blue. Such a belief is true or false depending on the circumstances. So one can combine the view that there are color beliefs with the view that colors are nonrobust properties as well as the view the colors are robust properties.

What if view 1, the theory that moral decision-making consists in having desires alone, is correct? Many philosophers who have been inclined toward such a view have had difficulty explaining the apparent belief-like nature of moral commitments. It seems quite intuitive to describe a person who is against the Iraq war as a person who believes the Iraq war is wrong. For this reason, philosophers such as Simon Blackburn have acknowledged that moral commitments, while essentially desires, often have belief-like characteristics. Such states are "quasi-beliefs" or "pseudo-beliefs": While they are not genuine beliefs, they are sufficiently belief-like to play a similar role.

If one is sufficiently motivated by the Humean considerations cited above to hold view 1, and deny the possibility of any moral beliefs, then one can appeal to such a notion of "quasi-belief" to articulate a range of

positions on moral truth and moral metaphysics. Perhaps, as A.J. Ayer claims, moral commitments are not sufficiently belief-like to have propositional contents, hence there are no moral truths. Perhaps, as Blackburn claims, moral commitments are similar enough to beliefs that we can consider such commitments capable of being assessed as true or false.

In a similar vein, if a philosopher who denies the existence of genuine moral beliefs holds that moral commitments are nonetheless "quasi-beliefs," such a philosopher can take a range of positions on whether such "quasi-beliefs" denote properties. Perhaps these "quasi-beliefs" denote no properties whatsoever, or they may denote robust properties, or they may denote non-robust properties. There are a range of options here, and the position one takes regarding these options is not settled simply by claiming that moral commitments consist of desires alone. To cite three examples of philosophers who claim that moral commitments consist of desires alone yet hold different views on moral metaphysics: Ayer holds that there are no moral properties; Blackburn's (2001) view seems best characterized that there are moral properties but no robust moral properties; and Copp's "realist expressivism" involves endorsing the view that there are

robust moral properties. There are a wide range of combinations of positions on moral psychology, metaphysics, and truth, and insufficient thought has been given to the array of positions in logical space.

The Cognitivism/Noncognitivism Distinction

In light of the distinctions made above, perhaps we can answer the following questions: What is "cognitivism"? What is "noncognitivism"? It is quite typical to find discussions of meta-ethical theories that are centered on the question of whether a certain view is a "cognitivist" or "noncognitivist" view. Thus in his recent book *Thinking How to Live*, Alan Gibbard asks the question "Is my theory, then, a form of non-cognitivism?" (Gibbard 2003, 183). After considering his various commitments on semantic, psychological, and metaphysical matters, he offers the following answer:

> In one sense, then, I am a "non-cognitivist": I draw from central aspects of that tradition. This may not lead me, though, to anything that staunch "cognitivists" deny (ibid., 184).

Such an answer might not satisfy one who thought that the question of whether a view is a "noncognitivist" one was a clear, well-defined question in the first place. As I will argue in this chapter, there is no such clear

question of whether a view is "noncognitivist," and Gibbard has good reason to give such a noncommital answer to this sort of question. If there were a clear and simple distinction between "cognitivism" and "noncognitivism," then there would be room for a clearer answer. In preparation for the discussion of meta-ethical theories that will follow in the next two chapters, it will be important to consider this purported distinction, and what role it has played in the philosophical literature.

In a recent entry in the *Routledge Encyclopedia of Philosophy* on "Analytic Ethics," Peter Railton offers the following definitions of "cognitivism":

> C1. Cognitivism is defined as a position on the truth of moral statements: "Are [moral statements] factual statements capable of being literally true or false?" (Railton 1998, 220)
>
> C2. Cognitivism is defined as an ontological position, committed to substantial moral facts, either naturalistic or nonnaturalistic: "Cognitivists focus on whether the facts to which they claim moral judgments correspond are discovered from experience, or whether they occupy a different realm, as do mathematical facts" (ibid.)

Within the same encyclopedia, Garrett Cullity presents a third, distinct definition of cognitivism:

> C3. Cognitivism is defined as a psychological position, according to which moral commitments are beliefs: "Thus cognitivism about the state of moral judgment—the view that it consists essentially in belief…" (Cullity, 512).

Corresponding to these three definitions of "cognitivism," it is quite easy to discover the definition of "noncognitivism." "Noncognitivism" is the conjunction of the negation of C1, the negation of C2, and the negation of C3. Thus, turning again to the encyclopedia entries by Railton and Cullity, we discover:

> NC1: Noncognitivism is defined as a position on the truth of moral statements. Moral statements are not capable of truth; They are "capable only of greater or lesser appropriateness or efficacy…" (Railton 1998, 220).
>
> NC2: Noncognitivism is defined as a position that does not involve commitment to moral ontology. The appeal of noncognitivism is its "freedom from metaphysical baggage" (ibid., 221)
>
> NC3: Noncognitivism is defined as a psychological position, according to which moral judgment does not

> consist in belief. Noncognitivism is the "denial of cognitivism" according to which "moral judgment…consists essentially in belief" (Cullity, 512)

It would not be difficult to find, while searching through the philosophical literature, a number of definitions of "cognitivism" in terms of some combination of C1, C2, and C3, and definitions of "noncognitivism" in terms of NC1, NC2, and NC3.

The trouble with these definitions is that they collapse several distinct issues—issues concerning truth (C1, NC1), metaphysics (C2, NC2), and psychology (C3, NC3)—into the single issue of "cognitivism" vs. "noncognitivism". The great deal of confusion in standard discussion of the matter of "noncognitivism" results from the fact that, as defined above, "noncognitivism" is not a single view, but a set of views. In the earlier sections of this chapter, I distinguished the issues concerning truth, metaphysics, and psychology in order to show that the "cognitivism"/ "noncognitivism" distinction presents a misleading view of how many possible meta-ethical positions there are in logical space. Given that it was possible to describe the various positions on truth, metaphysics, and psychology without making essential appeal to any of the

others, we have a clear notion of how many conceptually distinct positions on meta-ethics there are in fact. I will characterize a mistake commonly made due to the assumption that there are only two such positions.

The "Noncognitivism" Fallacy

There is a quite simple logical fallacy at the root of a great deal of confusion over this distinction. I will call this particular instance the "noncognitivism" fallacy. What is at root of the "noncognitivism" fallacy is that "noncognitivism" is supposed to be a complement class term. When one uses a complement class term properly, one uses it to pick out everything in a certain domain that is not a member of a particular class. For instance, when one wants to talk about all of the persons in the United States who are not members of the Republican party, one can use the term non-Republicans. This is useful, for it helps one to avoid writing out a long list of those persons who are not Republicans, i.e. Democrats, Greens, Right-to-Lifers, Libertarians, etc. There are a number of legitimate uses of complement class terms in contexts similar to those in which "cognitivism" and "noncognitivism" are commonly discussed. Take, for example, the distinction between naturalistic and nonnaturalistic theories. It is quite clear that a theory must be either naturalistic or

nonnaturalistic. Roughly, either a theory adheres to the methods that are consistent with science, or a theory fails to adhere to such methods.

Given that the issues discussed in the three sections above are distinct, it is quite clear that the definitions of "cognitivism" and "noncognitivism" given above are descriptions not of a single position, but of a conjunction of conceptually distinct positions. Thus "cognitivism" is defined as C1&C2&C3, whereas "noncognitivism" is defined as NC1&NC2&NC3. The source of what I am calling the "noncognitivism" fallacy is the assumption that the denial of "cognitivism" will lead one to "noncognitivism," and vice versa. However, once it is made clear that each of these positions as described in the philosophical literature is not a single view but rather a conjunction of distinct views, then one realizes the mistake. The problem is that one need only reject one conjunct in order to reject a conjunction. One can describe a position such as the following: the position that moral utterances are straightforwardly true or false, there are no robust moral properties, and moral decisions and commitments contain both beliefs and desires. Is this "cognitivism" or "noncognitivism"? How are we to decide? The distinction does not seem to offer room for such a position. Yet, as I

have argued above, such a position is clearly conceivable due to the fact that the issues of moral metaphysics, moral psychology, and moral truth can be addressed separately.

In the following two chapters, I will discuss the ramifications of these distinctions on the proper understanding of two major meta-ethical theories, expressivism and constructivism. Recent discussion of such views has been clouded by the confused views regarding the range of meta-ethical possibilities that I have discussed above. When one carefully distinguishes the relevant issues, it is possible to see what is and what is not required to articulate such a position.

CHAPTER 6

ON THE PURPORTED INCOMPATBILITY OF EXPRESSIVISM AND DEFLATIONISM

There has been a great deal of discussion in the recent philosophical literature of the relationship between deflationary, particularly minimalist, theories of truth and the expressivist meta-ethical theory. One group of philosophers contends that deflationism and expressivism are compatible, the other group contends that such theories are incompatible. Following Simon Blackburn (manuscript), I will call the former position 'compatibilism' and the latter position 'incompatiblism.' Even those compatibilist philosophers who hold that there is no conflict or tension between these two theories—deflationism and expressivism—typically think, as Field (1994) and Jackson, Oppy, and Smith (1994) do, that some revision of deflationism is required to accommodate expressivism.

In this chapter, I argue against the philosophers who hold that expressivism and deflationism are incompatible as well as those philosophers who hold that expressivism requires a revision of deflationism. In the first section, I argue that there is no incompatibility between

deflationism and expressivism. The claim that there is such an incompatibility is based on a misunderstanding of the historical roots of expressivism, the motivations behind the expressivist theory, and the essential commitments of expressivism. I present an account of the expressivist theory that is clearly consistent with deflationism. As I have done throughout the dissertation, where it is possible I will focus my argument on the minimalist account of truth, with the intention of formulating an argument that can generalize to other deflationary accounts of truth.

It is important to note that this is not simply a verbal dispute regarding the proper uses of the terms 'deflationism' and 'expressivism.' Any such dispute would be of little theoretical interest. The concern that I will address in this chapter is a substantive one. Opponents of the compatibilist position have incorrectly overlooked a possible philosophical position. Regardless of the labels applied to such philosophical positions, the very possibility of these positions being correct ought not to be denied without sufficient argument and consideration of the complete array of meta-ethical views in logical space.

In the following section, I argue that the aforementioned attempts to revise deflationism in order to

accommodate expressivism are unsuccessful as well as unnecessary, and such revisions lead to serious difficulties for expressivism and deflationism.

In the final section of the chapter, I consider the bearing of the Frege-Geach problem upon the issues discussed earlier in the chapter. I argue that the existence of such a problem is not relevant to the issue of whether or not expressivism and deflationism are compatible. Furthermore, in light of the account developed by Horwich (1993) and defended in Chapter 4 of this dissertation, I defend Horwich's claim that minimalism can help resolve Frege-Geach problem (*pace* Dreier 1996 and Gibbard 2003b).

Are Minimalism and Expressivism Compatible?

Paul Boghossian (1990), Crispin Wright (1985, 1992, 1996a), and John Divers and Alexander Miller (1994) have argued that the minimalist account of truth is incompatible with the expressivist theory of morality, whose proponents traditionally have held that moral utterances are neither true nor false. This purported incompatibility between minimalism and expressivism rears its head not only for the minimalist, but also for other deflationists regarding truth. As I explained in Chapter 2, all of these various deflationist accounts of truth, while differing from

minimalism in the details of their theories of the notion of truth, have in common a viewpoint that presents a quick route to the attribution of truth to moral utterances.

The apparent difficulty here results from the view that the expressivist, in order to distinguish her view from those who hold distinct meta-ethical positions, must hold what I have called the Denial Thesis, according to which moral utterances are neither true nor false. Given that the Denial Thesis has been held by philosophers whose views were forbears of expressivism, such as Hume and Ayer, as well as recent expressivist writers such as Gibbard (1990), it is widely believed that the Denial Thesis is an essential element of any legitimate expressivist view.

The apparent trouble with squaring expressivism with deflationism is that there is only a trivial difference, on deflationary theories of truth, between the assertion that p and the assertion that it is true that p. I will reiterate the reasons that this is so for the minimalist theory. The reasons why this is so were discussed in detail in Chapter 2, "On Considering Moral Utterances True or False." In discussions of moral matters, it is quite clear that utterances that have the appearance of genuine assertions are made, and such "assertions" would seem to express true or false propositions regarding moral issues.

A person may sincerely "assert" "Great inequalities in the distribution of wealth are unjust," and such an "assertion" would seem to express the belief that great inequalities in the distribution of wealth are unjust.

However, given that the object of this apparent belief is the proposition that great inequalities in the distribution of wealth are unjust, the minimalist would have to accept that such a proposition is capable of being straightforwardly true or false, *pace* the traditional expressivist view. It may still be possible, as I discussed in Chapter 2, for a minimalist to deny that there are in fact propositions expressed by these apparent assertions of moral claims. As I have argued previously, the most coherent position for a minimalist to hold is one according to which there are moral propositions, and thus moral utterances are capable of being true or false. On this view, the following would be a legitimate instance of the minimalist schema:

> I: <Great inequalities in the distribution of wealth are unjust> is true iff great inequalities in the distribution of wealth are unjust.

Thus, according to a minimalist account of truth, given the further assumption that moral utterances genuinely express propositions, such utterances must be either true or false.

Such an admission is tantamount to the rejection of traditional expressivism.

In "The Status of Content," Boghossian presents a version of this incompatibilist argument directed toward the work of A.J. Ayer. Boghossian summarizes Ayer's account of the redundancy theory of truth (an account described above in Chapter 1, in the section entitled **The Redundancy Theory**) and his expressivist account of ethics, and claims that Ayer failed to recognize

> the *tension* between such a [redundancy theory] conception [of truth] and a non-factualist thesis about a given range of assertoric discourse (Boghossian 1990a, 163).

The particular "non-factualist thesis" Boghossian has in mind here is Ayer's expressivist account of moral discourse.

Boghossian notes, for reasons similar to those cited in Chapter 2 of this dissertation, that moral utterances fit many of the criteria for straightforward assertoric discourse of the sort to which one would attribute truth. Moral utterances are meaningful, declarative, and are embedded in sentences formed by truth functional connectives (conjunction, disjunction, negation, conditionals) and within propositional attitude contexts.

There is at least a *prima facie* reason to regard such utterances as being on a par with nonmoral utterances.

Given that we have such good reason to attribute truth to moral utterances, Boghossian contends that Ayer's redundancy theory and expressivism are in conflict. In order to resolve the conflict, one needs a more "robust" theory of truth that will allow one to distinguish moral discourse, which is not straightforwardly true or false, from other realms of discourse that can be straightforwardly true or false (such as, for example, physics). Boghossian writes:

> A non-factualism about any subject matter requires a conception of truth stronger than the deflationary: it is committed to holding that "true" stands for some real, language-independent property, eligibility for which will not be certified solely by the fact that a sentence is declarative or significant (Boghossian 1990a, 165.

It is important to note here that this argument is framed in response to Ayer's views.

The argument presented by Wright differs from that of Boghossian insofar as it is clearly directed at contemporary deflationists. Wright argues that any expressivist or projectivist account that allows for the

attribution of truth to moral utterances will collapse into "moral realism." In his review of *Spreading the Word*, Wright makes this charge against Blackburn's quasi-realism. While not explicitly mentioning deflationism or minimalism, Wright notes that Blackburn's account seems to pave the way for the attribution of truth to moral statements, a result that I have argued (in Chapter 2) that deflationists and minimalists ought to accept. Thus, in the terminology of Chapter 2, Wright is worried about the compatibility of the Attribution Thesis and expressivism:

> If it really can be explained—and Blackburn takes imaginative strides toward doing so—how the moral projectivist can acknowledge the susceptibility of moral judgements to conditional and other forms of embedding, and even how we can have a worthwhile truth predicate for them, then, so far from vindicating a form of moral-anti-realism, why has it not been explained how the moral realist can, in effect, cut past the epistemological difficulties which beset non-naturalism without incurring any obligations to furnish reductions? (Wright 1985, 318)

I will call this argument of Wright the "collapse argument." A more recent version of the "collapse

argument" is clearly directed at minimalism regarding truth[37]:

> With each of these minimalisms in place, almost all of the areas which have traditionally provoked realist/anti-realist debate—ethics, aesthetics, intentional psychology, mathematics, theoretical science, and so on—will turn out to traffic in truth-evaluable contents, which moreover, when the disciplinary standards proper to the discourse are satisfied, we are going to be entitled to claim to be true. So two traditional anti-realist paradigms are immediately under pressure: *expressivism*—the denial that a target discourse, despite an apparently assertoric surface, really deals in truth-evaluable

[37] It is worth noting that Wright is using the term 'minimalism' to refer to a different theory from the minimalist theory of truth I have discussed throughout this dissertation. The characteristics of Wright's 'minimalism' and the minimalist theory of truth that lead to the problem discussed here are identical, though, and the subsequent philosophical literature has taken Wright to be criticizing not only 'minimalism' (in his sense) but all deflationist and minimalist accounts of truth. Also, Wright claims that 'minimalism' (in his sense) is what all deflationary theories essentially amount to, or ought to be: "…minimalism about truth, as described in this and the succeeding chapter…is just what the deflationist trend comes to (what would-be deflationists like Horwich *ought* to advocate)" (Wright 1992, 12). For this reason, it will not be necessary here to go into great detail regarding Wright's 'minimalism' and the differences between it and deflationary accounts.

contents—is not going to be an option… (Wright 1996a, 864)

The collapse argument can be paraphrased without using the somewhat vexed jargon Wright uses. The point Wright is making is as follows:

1. In order for expressivism to be a distinct philosophical position, there must be a significant difference between expressivism and moral realist positions.
2. The only significant difference there could be between expressivism and moral realism is the different approaches taken by the expressivist and the moral realist to the attribution of truth to moral utterances.
3. If Blackburn's account, or a deflationist or minimalist account, is correct, then the expressivist ought to attribute truth to moral utterances in the same fashion as the moral realist.
4. Thus, expressivism is not a distinct philosophical position.

The problematic premise in this argument is premise 2. As I have noted in Chapter 4, there is at least one significant difference between the expressivist and the

moral realist, namely that moral realists hold that there are robust moral properties and the expressivist does not. Thus it would not be correct to hold, as Wright does, that the only significant matter of dispute between the expressivist and the moral realist is over the attribution of truth to moral statements. Perhaps, however, implicit in Wright's argument is the claim also made by Boghossian, that expressivism's commitment to a certain view of truth renders it incompatible with deflationism. In the quote from Wright 1996a, it is worth noting that Wright presents as the *entire* definition of expressivism the claim that moral discourse does not "really deal in truth-evaluable contents."

The position presented by Boghossian and Wright, according to which deflationism and expressivism are incompatible, is based on a particular account of expressivism. This account reflects the commitments of those philosophers who first articulated an expressivist view, such as Hume and Ayer. It is not often noted that these philosophers had accounts of truth and meaning that led them to conclude that the expressivist should refrain from claiming that moral utterances are straightforwardly true or false. As I will discuss in the following two sections, Hume drew this conclusion based on the

conjunction of his views on morality and a correspondence theory of truth, and Ayer drew this conclusion based on the conjunction of his views on morality, a redundancy theory of truth, and a verificationist semantics. The minimalist who rejects the correspondence theory of truth and the verificationist theory of meaning need not draw the same conclusions as Hume and Ayer[38].

Hume, Correspondence Truth, and the Denial Thesis

The meta-ethical theory on which I have focused in this chapter, expressivism, was, as proponents of the theory acknowledge, first formulated in David Hume's great work *A Treatise of Human Nature*. In the *Treatise*, Hume states quite clearly the view that utterances in moral discourse are neither true nor false. This is the original source of the view that expressivism is commited to the Denial Thesis. However, unlike more recent philosophers, Hume makes it quite clear that this position on moral discourse is a consequence of two distinct theories: a metaphysical position on the nature of moral judgments, and a particular semantic theory of truth, namely a clear and simple version of the correspondence theory.

[38] Devitt (1996) makes essentially the same point regarding Ayer. I am greatly indebted to Devitt's discussion of these matters.

In the *Treatise*, Hume accounts for moral judgments in terms of the passions, or emotions. Only a passion can move a person to act, according to Hume, and so any moral decision-making must consist in having certain passions. Such passions are dependent upon the mind of the individual making the moral decision. If truth or falsehood consists of agreement between representations in the mind and mind-independent objects, as it does on Hume's formulation of the correspondence theory, then it is clear that Hume must deny that moral thoughts and moral utterances are capable of being true or false.

Hume's argument for the claim that moral sentences cannot either be true or false may be summarized as follows:

P1: (Semantic premise) Truth or falsehood consists in agreement or disagreement of sentences with mind-independent entities.

As Hume states the theory, "contradiction" of truth "consists in the disagreement of ideas, consider'd as copies, with those objects, which they represent" (Hume 2000, 267). Such contradiction clearly occurs in cases when one believes, for example, in a dream that one is sitting at a desk, but no desk corresponds to the idea one has, the mental representation of a desk.

P2: (Metaphysical premise) There are no mind-independent moral facts, but rather only passions and volitions in the mind.

Hume expresses this point as follows:

> A passion is an original existence, or, if you will, modification of existence; and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, sick, or more than five foot high (Hume, 2000, 266).

From P1 and P2 it follows that:

C: Moral thoughts and utterances are neither true nor false.

Hume state the conclusion as follows: "'Tis impossible, therefore, that this passion can be oppos'd by, or be contradictory to truth and reason" (ibid., 266-67).

With the original argument from expressivism to the Denial Thesis in hand, it ought to be clear that Hume's argument is not sound if we reject a correspondence theory of truth. If we reject premise P1, the correspondence theory, in favor of a minimalist account of truth, as I have argued we should in Chapter 1, it is clear that the

conclusion C, the Denial Thesis, does not follow from Hume's account. Philosophers have overlooked the fact that the Denial Thesis follows from expressivism on the Humean argument only on the assumption that the correspondence theory of truth is correct.

Ayer, Verificationism, and the Denial Thesis

As I have discussed in detail above, in the section **Are Minimalism and Expressivism Compatible?**, one of the influential arguments for incompatibilism is presented in response to the views of Ayer. Paul Boghossian contends that Ayer failed to recognize the "tension" between the redundancy theory of truth and an expressivist account of ethics (Boghossian 1990a, 163). Given that this debate has been framed in response to Ayer's views, it is not solely of historical importance to investigate whether there was in fact a tension in the views Ayer presented in *Language, Truth, and Logic*. In fact, when Gibbard describes his view in *Wise Choices, Apt Feelings* as a "non-cognitivistic" one due that "according to it, to call a thing rational is not to state a matter of fact, truly or falsely," he cites in a footnote his debt to Ayer's "non-cognitivistic treatment of moral language…" (Gibbard 1990, 8).

A closer inspection of Ayer's account reveals that Boghossian is incorrect regarding the supposed tension in

Ayer: The views presented in *Language, Truth, and Logic* are fully consistent, and Ayer's case for the Denial Thesis rests on a view that deflationists need not accept, a verificationist account of meaning. Ayer formulates his version of the redundancy theory of truth in terms of propositions. Propositions are the bearers of truth and falsity. Ayer writes:

> …to say that a belief, or a statement, or a judgement, is true is always an elliptical way of ascribing truth to a proposition, which is believed, stated, or judged.

That this is an important aspect of Ayer's view will become clear when we consider his view on whether moral utterances involve propositional content.

As noted above in Chapter 1, on Ayer's account, the addition of the words 'is true' to an assertion does not change the content of the original assertion. These words are redundant. "When one says that the proposition 'Queen Anne is dead' is true, all that one is saying is that Queen Anne is dead" (Ayer 1936, 88). As I also noted in Chapter 1, there are many disadvantages to such a redundancy theory that are not shared by other deflationary accounts such as minimalism and disquotationalism.

What has not been noted sufficiently is that in his discussion of moral language, Ayer makes a quite strong, rather implausible claim regarding moral judgments. Ayer notes a difficulty with accounts of morality which make appeal to rational intuition as the basis for moral knowledge. If we take such an intuitionistic view seriously, how do we determine which purported moral intuitions are the correct ones? If one person claims to know by intuition that sacrificing the life of one person to save the lives of five is right, and another person claims to know by intuition that sacrificing the life of one person to save the lives of five is wrong, how do we decide which intuition is the correct one? It seems we have no way of verifying the appeals to intuition involved in ethical judgment.

In the discussion of Ayer by Boghossian and the citation of Ayer by Gibbard, it is not noted that in addition to a redundancy theory of truth, Ayer held a verificationist account of meaning. On Ayer's account, an expression is meaningful only if it can be verified. Any expression that is not verifiable does not express a proposition, but rather a mere "pseudoproposition" that not capable of being true or false. It is due to Ayer's view that moral judgments are unverifiable—hence they express

"pseudopropositions"—that Ayer held that moral utterances are incapable of being true or false. Boghossian overlooks this verificationist constraint—With such a constraint in place, it is clear that Ayer's views on morality and truth are consistent.

As was the case with Hume's argument for the denial thesis, it is very important to note that the argument in Ayer from expressivism to denying that moral utterances are capable of being true or false rests on assumptions the contemporary deflationist need not share. A minimalist about truth who holds a use-theoretical account of meaning of the kind described in Chapter 3 of this dissertation clearly would reject a crucial step in Ayer's argument. Not being a verificationist, the use-theorist can reject Ayer's claim that moral utterances express only pseudopropositions. Thus neither of the historical arguments from expressivism to the Denial Thesis ought to move a philosopher convinced of expressivism and deflationism to hold that moral utterances are incapable of being true or false.

Minimalism, Expressivism, and the Attribution of Truth

The view that moral utterances are not straightforwardly true is a theory that is not required for a formulation of an expressivist theory. As I have shown

in the sections above, Hume and Ayer drew the conclusion that moral utterances are neither true nor false from their commitment to, for Hume, a correspondence theory of truth, and for Ayer, a verificationist theory of meaning. A philosopher who holds neither of these theories will naturally draw different conclusions regarding the truth and falsehood of certain utterances.

Regardless of the failure of the traditional expressivist arguments to motivate deflationists to accept the Denial Thesis, one can still ask whether this thesis is nonetheless an essential element of the expressivist theory. In order to see whether this is the case we must look to the arguments that have motivated expressivism as well as the details of the theory. If neither the motivations behind expressivism nor the details of the theory is in conflict with deflationism, then it is clear that the case for incompatibilism is a failure.

The crucial motivations behind expressivist views, historically, are as follows. First, expressivists are motivated by a view on the metaphysics of morality[39].

[39] Insufficient attention has been paid to the centrality of metaphysical concerns in the arguments for expressivism. One notable exception is Devitt 1996. While my debt to this paper should be obvious, I disagree in part with the characterization of 'nonfactualism' given by Devitt. According to Devitt, theories that are 'nonfactualist': (a)

Expressivists have denied the existence of robust moral properties for reasons of the sort discussed above in Chapter 5. In particular, expressivists are dubious of the claim that there could be moral properties that could play a significant causal-explanatory role. The other crucial motivation behind expressivism is that expressivists want an account of moral language that can clearly explain the close tie between moral commitments and motivation. This aspect of expressivism has a historical lineage beginning with the work of Hume cited above, particularly Hume's claim that the passions alone can motivate, and reason cannot. Due to this account of motivation, expressivists have claimed that moral utterances serve the purpose of expressing some mental state that plays a motivating role, such as an emotional reaction, a desire, acceptance of a norm, or what Gibbard (2003b) calls planning.

The minimalist account of truth shows the possibility of a coherent position that is consistent with these motivations behind expressivism while allowing for the

---

"deny the existence of a reality with a nature to be explained" (Devitt 1996, 171); (b) cannot provide any explanation of the purported reality. While I attempt to spell out a view similar to (a) in this chapter, I do not think (b) characterizes nonfactualism well. To cite one example, a prominent nonfactualist account of morality, Gibbard's *Wise Choices, Apt Feelings*, contains a lengthy and sophisticated explanation of morality based on evolutionary considerations.

possibility that moral utterances are capable of being true or false. As I have discussed in great detail in Chapters 4 and 5 of this dissertation, it is possible for the expressivist to hold the view that there are no robust moral properties and facts as well as the view that moral utterances serve to express desires (or similar mental states) independently of any commitments on a particular account of truth. As I argued in Chapter 4, there is no inconsistency between holding that moral utterances are straightforwardly true and holding that there are no robust properties picked out by moral predicates such as 'right,' 'just,' 'good,' and 'evil.' As I argued in the **Moral Psychology** section of Chapter 5, the issue of whether moral commitment consists of beliefs only, desires only, or some combination of the two is independent both of this metaphysical issue and the matter of whether truth ought to be attributed to moral utterances.

Even if the point I have made above is correct, and the motivations behind expressivism are not in conflict with minimalism regarding truth, one might still worry that the elements of the theory are conflict with minimalism. Gibbard offers a clear definition of expressivism in *Wise Choices, Apt Feelings*. Gibbard writes:

> According to any expressivistic analysis, to call something rational is not, in the strict sense, to attribute a property to it. It is to do something else: express a state of mind. It is, I am proposing, to express one's acceptance of norms that permit the thing in question (Gibbard 1990, 9).

Gibbard's claim that expressivism involves the expression of a state of mind, the acceptance of certain norms, is not in conflict with minimalism or the Attribution Thesis. A theorist who accepts this could also accept that moral utterances are capable of being straightforwardly true or false.

There might, however, be a worry regarding Gibbard's claim that calling something rational does not involve "attributing a property to it." It might seem that this aspect of the definition of expressivism is in conflict with the position I have defended in this dissertation. As I argued in Chapter 4, the best way to characterize the metaphysical position implicit in expressivism is not to deny that moral utterances involve the attribution of properties, but rather the deny that such utterances involve the attribution of robust properties. In the final section of this chapter, **The Embedding Problem**, I present a further reason the expressivist ought to accept this

position of moral properties, namely that it resolves the Frege-Geach embedding problem.

A closer look at what Gibbard says regarding the metaphysics of expressivism makes it clear that this is only an apparent conflict, not a genuine one. As Horwich notes in "Gibbard's Theory of Norms," in order to understand Gibbard's denial of the existence of properties and facts, we need to keep in mind that Gibbard has a "certain 'thick' conception of 'fact'" in mind" (Horwich 1993, 70). Gibbard has a thick conception of properties in mind as well. On the minimalist account of properties presented in Chapter 4, there is only a trivial difference between saying that x is F and saying that x has the property of being F. On the related minimalist account of facts, there is only a trivial difference between asserting that p and asserting that it is a fact that p.

It is clear that Gibbard need not deny that, in this minimal sense of properties and facts, we attribute moral properties to acts and persons, and moral utterances are assertions of moral facts. On the other hand, Gibbard clearly denies that moral language involves the attribution of a "thicker" robust property, a property that plays a causal-explanatory role. Gibbard denies that there are robust properties in the former sense in *Wise Choices, Apt*

Feelings when he claims that "Normative talk is part of nature, but it does not describe nature" (Gibbard 1990, 7).

In his more recent work, as I discussed in Chapter 4, "Minimalist Schemas: Truth, Properties, and Facts," Gibbard draws on the distinction between concepts and properties. Certain concepts are naturalistic concepts, concepts "that arise in strict empirical science and in everyday causal explanations of our experience and observations…" (Gibbard 2003b, 32). Moral concepts are non-naturalistic concepts—Moral concepts are concepts that do not figure into causal laws, causal explanations, and observations. Thus, the property we are thinking about when we have thoughts involving moral concepts is, while a property that belongs to the natural world, not itself a property that figures into causal explanations. In the sense of 'robust' I described in chapter 4, moral properties are not robust properties on Gibbard's view. Given that Gibbard is expressing a denial of the existence of robust moral properties, it seems that what is required to fit Gibbard's definition of expressivism is an account that does not requiring postulating robust moral properties. For reasons presented in Chapter 4, the conflict between expressivism and this account of moral properties is only an apparent one, not a genuine one.

Despite the fact that no convincing argument has yet been given for incompatibilism, incompatibilist arguments have been very influential. In fact, many of those philosophers who accept compatibilism have been moved by incompatibilist considerations to revise deflationism or minimalism to accommodate expressivism. These attempts have been unsuccessful. This is a further reason to accept a compatibilist account of the kind on offer in this dissertation, an account according to which there is no need for a philosopher who accepts expressivism and minimalism to revise either theory. As I will discuss in the following two sections, **Minimalism Regarding Truth and Minimalism Regarding Truth Aptness** and **Field's Response to Incompatibilism**, the options offered to the minimalist are ill-motivated in light of the arguments presented above, and lead to several unacceptable consequences.

Minimalism Regarding Truth and Minimalism Regarding Truth Aptness

Michael Smith (1994a and 1994b) and Frank Jackson, Graham Oppy, and Michael Smith (1994) have argued that deflationism and expressivism are compatible by appealing to a distinction between sentences that are true, and sentences that are truth-apt. A sentence that fails to be truth apt is one that is "not in the business of being true

or false," i.e. such a sentence does not express a proposition (Jackson, Oppy, and Smith 1994 287). According to their account, the expressivist ought to claim that moral utterances are similar to sentences such as 'Close the door!' or 'Hello' that fail to express a proposition, and thus are not candidates for truth values. Therefore an expressivist should hold that moral statements such as "Great inequalities in the distribution of wealth are unjust" serve only to express the desire for great inequalities in the distribution of wealth to be redressed, without expressing a proposition. Saying 'Great inequalities in the distribution of wealth are unjust' would be equivalent to saying 'Redress inequalities in the distribution of wealth!' which is a sentence that is no more capable of being true or false than "Close the door!"

In this respect, Jackson, Oppy, and Smith are suggesting that the contemporary deflationist-expressivist ought to embrace an account of the semantics of moral utterances not unlike that of Ayer's. Such an account would require a substantive theory to distinguish truth-apt utterances from those utterances that are not truth-apt. They do not propose endorsing a verificationist account along Ayer's lines. Jackson, Oppy, and Smith suggest that one can appeal to an account of propositions tied closely

to psychology in order to make the truth-apt/non-truth-apt distinction: Only utterances that express beliefs are truth-apt. Utterances that express desires are not truth-apt.

There are two major difficulties with the Jackson, Oppy, and Smith resolution of the supposed incompatibility between expressivism and deflationism. The first is that it is inconsistent with the goals of minimalist accounts to appeal to a distinction between truth-aptitude and truth in order to deny that utterances that seem *prima facie* to be expressions of truths do in fact express truths. Minimalists are attempting to characterize the use of the predicate 'true' by ordinary speakers of a language, not a more complicated, philosophically loaded notion of truth that disagrees with common practice. Unlike a sentence such as 'Close the door!' a sentence ordinary speakers would regard as neither true nor false, the commonsense view of moral statements such as 'Killing is wrong' is that such sentences are either true or false, and hence truth-apt.

The second difficulty with the Jackson, Oppy, and Smith account is that such an account leaves only the possibility of a rather crude form of expressivism that fails to accord well with our ordinary practice of moral

argument. For example, moral sentences are usually assumed to be capable of serving as premises in valid arguments. If moral sentences are not truth-apt, one would have to reject the claim that the following argument is valid:

P1: If murdering innocent people is always wrong, then murdering a small group of innocent people to save the lives of a larger number of innocent people is wrong.

P2: Murdering innocent people is always wrong.

C: Murdering a small group of innocent people to save the lives of a larger number of innocent people is wrong.

For, if P2 and the antecedent and consequent of the conditional in P1 are not truth apt, one could not appeal to the truth of these claims and the form of the argument to explain why the truth of these premises would lead, necessarily, to the truth of the conclusion. The Jackson, Oppy, and Smith account resolves the incompatibilist worry by throwing the baby out with the bathwater.

Field's Response to Incompatibilism

In his paper "Disquotational Truth and Factually Defective Discourse," Hartry Field presents a similar response to the incompatibilist argument. Field offers an account of normative discourse that is supposed to accord with the traditional expressivist view that normative

utterances are not capable of being straightforwardly true or false. Field rejects the approach taken by Horwich (1998b), an approach that I have defended in this dissertation. In *Truth*, Horwich argues, that (as Field puts it) "a deflationist should adhere to the [minimalist truth] schema in the case of ethical sentences" (Field 2001, 243). Field contends that Horwich's account of such a view, combining minimalism and expressivism, fails to make it obvious "wherein the non-factualism lies" (ibid.).

It is not entirely clear why one could not characterize a view as "non-factualist" while adhering to the claim that ethical sentences express propositions and hence are legitimate instances of the minimalist schema. Field describes non-factualism as the view that "statements" related to non-factual matters "lack truth value…" (Field 2001, 238). Field then rejects such a definition of non-factualism as "superficial," for he acknowledges that "it is by no means obvious that a deflationist can't find a way to make sense of a use of 'true' according to which such statements *aren't* true or false" (ibid.). Field then claims that the truly interesting worry concerning non-factualism is "that the deflationist simply can't make sense of the distinction

between discourse that is fully fact-stating and discourse that isn't" (ibid., 238).

In light of the account of moral facts and properties spelled out in this dissertation, the charge that an account such as Horwich's could not make it clear "wherein the non-factualism lies" could be answered by reference to an account of the kind spelled out in Chapter 4. Expressivist philosophers, in effect, argued for their accounts by appealing to the metaphysical claim that there are no robust moral properties, and that there are no robust moral facts. Perhaps this account of metaphysics would provide sufficient "non-factualism" to answer this worry. That the expressivist would deny the existence of robust moral facts and properties marks a significant difference between her account and the accounts of "moral realist" opponents.

Given that Field raises in his paper issues related to whether deflationists generally can hold views of moral discourse such as expressivism, it is reasonable to suspect that Field is raising the incompatibilist worry. That this is so is implicit in Field's account of moral truth. As I have noted, Field is attempting to revise a deflationary theory of truth in order to answer the arguments for incompatibilism.

Field presents his view of the truth of moral utterances in terms of a modified version of Gibbard's (1990) adaptation of possible world semantics[40]. On Field's view, moral utterances are not straightforwardly true or false, but only true relative to a set of norms N. Thus we would say that "Great inequalities in the distribution of wealth are unjust" may be true or false, but only relative to the acceptance of a certain norm regarding just distribution of wealth in a certain world w. This relativity of truth of moral utterances to norms is made explicit in the formulation of instances of the deflationary schema—in Field's account, the disquotational schema—containing such utterances. In Field's example:

> (For any norm N) 'We should intervene in Bosnia' is true relative to N iff we should intervene in Bosnia relative to N (Field 1994 439).

The question as to whether moral utterances are absolutely true or false can be cast as the question of whether there is the need to relativize the truth of moral utterances to certain norms. Field's moral "factualist," who holds that there are straightforwardly true moral utterances, denies that such a relativization is required.

---

[40] Gibbard's original account will be presented in detail below in the section entitled **The Embedding Problem.**

There are several difficulties for Field's account of moral truth. This relativized account of moral truth does not seem to allow for serious moral commitments, and is at odds with the non-relativistic views generally held by expressivists. To see that this is so, it is important to look at the analogy Field raises between his "nonfactualist" account of moral discourse and a "nonfactualist" account of absolute time order. Field notes that a person fully aware of contemporary physics, and of the special theory of relativity, would realize that one cannot claim *simpliciter* that an event A occurred prior to event B. Any such claim must be made relative to a particular frame of reference. Hence, the instance of the disquotational schema that would be accepted by a person who is aware of such nuances of theoretical physics is the following, in which F specifies the frame of reference:

> 'A was earlier than B' is true relative to F if and only if A was earlier than B relative to F (Field 240).

So an individual who is fully aware of physics can only accept such a framework-relative instance of the truth schema, and would only accept the following as a "less explicit way of stating" the framework-relative schema instance (ibid., 241):

'A was earlier than B' is true iff A was earlier than B.

So, on Field's analogy, a person fully aware of the nature of moral utterances would only be able accept a truth schema instance such as R1 as a less explicit way of stating R2:

R1: 'Rape is wrong' is true iff rape is wrong.

R2: (For any norm N) 'Rape is wrong' is true relative to N iff rape is wrong relative to N.

Is it actually the case, as Field's analysis suggests, that whenever a person properly accepts a moral claim, such as the claim that rape is wrong, one only accepts it as true relative to one's own framework of norms? Wouldn't a person who takes such a moral commitment seriously want to claim that rape would be wrong regardless of the system of norms that she happens to accept? Moral commitments would seem to involve accepting claims that are independent from facts about one's own moral attitudes, rather than relative to such attitudes[41].

[41] I do not intend to suggest here that it will always be possible to find a determinate answer to moral questions. In cases where we happen to find two moral alternatives equally justified in light of what we know, and there is no way to determine that one of the alternatives is better, we ought to judge that each of these two alternatives is on a par with the other. For further discussion of the use of 'on a par' as a moral term, see Chang (2002). (I am

Prominent philosophical proponents of expressivism have stressed that expressivism ought not to be conflated with moral relativism, as Field's account seems to do. In his essay "Relativism and Nonrelativism in the Theory of Value," the expressivist Charles Stevenson presents an argument similar to the one presented above, contending that a relativist theory could not allow for moral commitment:

> A consistent relativist, when asked what is good or right, etc., would in effect discuss only what is or was considered right or good, etc., and thus would himself stand committed to no value judgments whatsoever. He would be a non-participant on evaluate issues—as no man, in practice, can be (Stevenson 1963, 71).

Contemporary expressivists such as Gibbard and Blackburn likewise stress that expressivism is not a relativist view. In a recent response to the suggestion of A.W. Moore that expressivists are committed to relativism, Simon Blackburn has argued that for an expressivist

> …there is nothing relativistic to say, along the lines that when people disagree with us rival views may be

---

grateful to Steve Ross for pressing an objection on this topic).

> ‘false for us’ and ‘true for them.’ A view we totally reject, such as the view that women deserve death for talking to men outside the family, is not ‘true for’ anybody at all, any more than it is true in Pakistan and false here. It is false, pure and simple (Blackburn 2002, 134).

The problem with squaring Field’s approach with the account of moral commitments offered by expressivists such as Blackburn, Gibbard, and Stevenson is that it does not allow for a claim to be a “pure and simple” truth or falsehood.

Third, Field’s view does not seem to allow for the possibility of moral disagreement. Typically, if one person were to claim that great inequalities in the distribution of wealth are just, and another were to claim that great inequalities in the distribution of wealth are unjust, we would think these two are contradicting each other. However, on Field’s analysis, there would be no such contradiction in the case where the different viewpoints are assessed relative to the acceptance of distinct sets of norms.

Fourth, Field does not provide sufficient motivation for his view that moral utterances function in a way analogous to indexical utterances. In the case of indexicals, speakers recognize that utterances of sentences

containing terms such as ‘today’ or ‘yesterday’ are true or false relative to a certain occasion. Recognition of this purported relativity of moral truths to norms is, as Field himself grants, not part of the ordinary speaker’s use of moral terms.

Can this Dispute Be Resolved?

It is possible that at this point in the debate, due to the fact that expressivists have traditionally assumed that moral utterances are incapable of being true or false, a philosopher could deny that the position I sketched above is in fact an expressivist position. One would have to ask at this point whether this is a substantial dispute regarding the motivations behind expressivist theories, and the details of the theory as spelled out by its proponents, or whether this is a purely verbal dispute about the proper use of the philosophical term ‘expressivism.’

Such a philosopher might present a convincing reason for thinking that the position I have sketched above is not an open option. If this were to happen, such an argument would be an argument about the substance of the issue, and would not be a purely verbal dispute. If, on the other hand, a philosopher does not give a reason to think that this position is not an open option, and merely insists on a certain restricted use of the philosophical term

'expressivism,' then this philosopher would be making a purely verbal point. If this were to happen, I would simply cede the term 'expressivism' to my opponent. My concern in this chapter has been to discuss the relationship between deflationary and minimalist theories of truth and a meta-ethical account of the kind proposed by Hume, Ayer, Stevenson, Gibbard, Blackburn, and other philosophers. If necessary, I could phrase the question posed in this chapter as the question of whether an account of the kind presented by these philosophers is consistent with deflationism and minimalism. Another approach, proposed by Horwich 1993, would be to suggest that an account of this kind is a revision of expressivism, a revision that has advantages over its traditional expressivist predecessors. Either strategy, it seems, would successfully answer this worry.

The Embedding Problem

There is an apparent problem for expressivist accounts of moral discourse. This problem is typically called the Frege-Geach problem, for it was pointed out by Peter Geach in his paper "Assertion," and Geach attributed the initial formulation of the problem to Frege. According to Geach, "Frege already made this point in his youthful work, *Begriffschrift*, many years before formulating his mature

views on sense and reference" (Geach 1965, 449). According to James Dreier, the problem was also independently discovered by Searle (1962). In order to avoid the issue of the proper attribution of this point, I will follow Dreier and call it the embedding problem. In this final section of the chapter, I will consider whether there is any reason to think that the embedding problem would create a special problem for the compatibilist, the philosopher who holds that deflationism and expressivism are compatible. I will argue that there is no reason to think that there is a special problem: First, the purported embedding problem is not an problem for theories of *truth*, but an issue regarding the proper account of the semantics of moral terms and the semantics of logical connectives such as the conditional; Second, in light of the account presented in Horwich 1993 and defended in this dissertation, the use theory of meaning and the minimalist theory of truth present an elegant solution of the embedding problem.

The embedding problem is a result of the combination of expressivist accounts of the meaning of moral utterances and a well-known feature of certain sentential contexts. On an expressivist account of moral discourse, as noted above, the central function of moral and normative claims is to express an attitude of a certain kind. Utterances

such as “Lying is wrong” and “Rape is abhorrent” express attitudes toward acts of lying and rape. One would not fully understand such utterances if one did not take account of the fact that the person who makes such utterances has a negative response to the relevant acts. In Gibbard’s terminology, the speaker who makes such utterances is expressing her acceptance of a certain norm: in the former case, the speaker accepts a norm prohibiting lying; in the latter case, the speaker accepts a norm prohibiting rape.

As the apt term ‘embedding problem’ suggests, difficulties are supposed to arise for the expressivist when moral and normative sentences are embedded within certain contexts. These are contexts where the embedded sentence is not asserted by the speaker uttering the containing sentence. To take an example not involving a moral expression, when one says ‘If it is raining heavily, then there will be puddles on 7th Avenue,’ one is asserting neither ‘It is raining heavily,’ nor is one asserting ‘There will be puddles on 7th Avenue.’

In a conditional statement, neither the antecedent nor the consequent is asserted. This is also the case with the sentences embedded in attributions of beliefs. When one

says 'Susan believes that it is raining outside,' one does not assert that it is raining outside.

There are, of course, contexts where sentences contained within a larger expression are all asserted. For example, in a conjunction, both conjuncts are asserted[42]:

A1: It is raining outside and there will be puddles on 7th Avenue.

Also, in attributions of knowledge, the sentence giving the content that is known is asserted:

A2: Susan knows that it is raining outside.

The apparent embedding problem for the expressivist results when normative and moral sentences appear in contexts where the contained sentences are clearly not asserted. Thus, take for example, L1:

L1: If lying is always wrong, then lying on your resume to get a better job is wrong.

The assertion of the entire sentence L1 does not require one to assert either that lying is always wrong or to assert that lying on your resume to get a better job is wrong. One could conceivably endorse this conditional without endorsing either the antecedent or the consequent.

[42] Blackburn apparently does not notice this point when he suggests that one needs to give a revisionary account of the meaning of 'and' in order to capture the meaning of conjunctions containing conjuncts that express attitudes (c.f. Blackburn 1984, 191-2).

If one can assert L1 without endorsing either of these statements, then it is clear that we cannot explain the meaning of the antecedent or consequent in terms of the expression of a disapproval. One can utter L1 without expressing disapproval of lying or disapproval of lying on one's resume to get a better job.

On the other hand, when one sincerely utters the sentence

L2: Lying is always wrong.

one clearly is endorsing the claim that lying is always wrong. Obviously, it is also the case that sincere utterances of L3 involves endorsing the claim

L3: Lying on your resume to get a better job is wrong.

If we take L1 and L2 as premises, and L3 as a conclusion then we have a valid argument, an instance of modus ponens:

Premise 1, L1: If lying is always wrong, then lying on your resume to get a better job is wrong.

Premise 2, L2: Lying is always wrong.

Conclusion, L3: Lying on your resume to get a better job is wrong.

What is troubling here is, as noted above, that the expressivist account of the meaning of moral statements would seem to require us to claim that the embedded

antecedent and consequent in L1 do not have the same meanings (respectively) as L2 and L3. For, if the meaning of a moral utterance consists in the expression of the acceptance of a norm, L2 and L3 involve such expressions of acceptance but the antecedent and the consequent in L1 do not. Thus it seems that the argument is of the invalid form

If p then q,

r,

therefore s

Rather than the valid form

If p then q

p

Therefore q.

Thus the apparent result is that the expressivist account turns what is intuitively a valid argument into a fallacy of equivocation.

In response to this objection, the expressivists Blackburn and Gibbard have offered revisionary accounts of the semantics of the conditional and the other logical connectives. The primary focus of these accounts has been the semantics of the conditional. On Blackburn's initial account in *Spreading the Word*, we are to understand conditionals containing moral sentences as expressions of

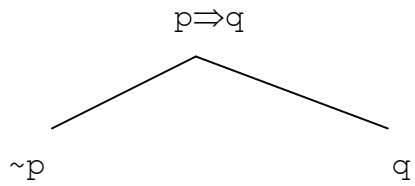
an attitude toward holding a combination of attitudes. Roughly, when I say that if lying is always wrong then it is wrong to lie on your resume to get a better job, I say that it would it would be irrational for a person who holds that lying is always wrong not to also hold that lying on your resume to get a better job is wrong. The attitude I express when asserting L1 is the attitude condemning a person who held L2 but failed to hold L3.

A number of philosophers[43] have argued that this account does not do justice to our sense that an argument from L1 and L2 to L3 is a valid argument. Instead, Blackburn's account seems to substitute for the claim that this argument is logically valid the claim that the morally right thing for a person who holds L1 and L2 to do is to hold L3. Rather than failing to make the correct deductive inference, a person who held L1 and L2 but denied L3 would manifest a moral flaw. Thus it seems that Blackburn solves the embedding problem at the heavy cost of denying one of the intuitions leading to the problem in the first place, the intuition that the argument from L1 and L2 to L3 is valid.

---

[43] Versions of this argument have been presented, as Blackburn (manuscript) notes, by Hale 1986, Wright 1987, Schueler 1988, Hurley 1989, Brighouse 1990, Zangwill 1992, Wedgwood 1997.

In response to this objection, Blackburn (1988a) presents a new account of the meaning of conditional statements. On this new account, Blackburn explains the conditional by appeal to a certain kind of logical diagram, a tree structure. A conditional statement of the form p⇒q[44] is true if either ~p or q[45]. Thus, we can depict the conditions under which p⇒q is true using the following diagram:



On Blackburn's new account, a person who endorses the conditional statement p⇒q is thereby committed to endorsing either ~p or q. In Blackburn's metaphor, this person is "tied to a tree," the tree structure above. If this person undertakes the commitment to p⇒q, and ties himself to the tree, then we can fairly say he is being inconsistent if he endorses p, but fails to endorse q. One is forced, in order to be consistent, to endorse the second branch of the tree if the first branch is rejected. It is thus clear

[44] I will use the following symbols for the logical connectives. If…then: ⇒. Not: ~. And: ∧. Or: ∨.
[45] For the sake of the discussion here, I will not try to settle whether this account of the truth tree of the conditional is correct.

that a person who endorsed L1 and L2 but did not endorse L3 would be inconsistent according to this account, and would demonstrate a flaw in reasoning by the lights of this account. It is still far from clear that Blackburn has presented an account that captures the *logical* inconsistency of such a person—Why is this not just a practical inconsistency? It is unclear that the notion of being "tied to a tree" resolves the worry raised regarding Blackburn's earlier account. There also remains some doubt as to whether this account can adequately explain the validity of all arguments containing embedded moral utterances (Wedgwood 1997, Kölbel 2002).

The approach to the embedding problem taken by Gibbard in *Wise Choices, Apt Feelings* draws on the resources of modal semantics in order to account for the meaning of expressions containing moral sentences and logical connectives. On a standard possible worlds account of semantics (such as Stalnaker 1976 or Lewis 1986), the content of a statement, the proposition expressed by the statement, is a set of possible worlds[46]. The content of

[46] Gibbard's discussion follows Stalnaker's account, rather than Lewis's, in construing a possible world as a "completely determinate way [one] thinks the world to be" (Gibbard 1990, 95). The Lewis/Stalnaker debate is outside of the scope of this dissertation.

'Grass is green' is the set of all possible worlds such that grass is green.

The meaning of the logical connectives is explained in terms of set-theoretic relations among sets of possible worlds. For example, the proposition expressed by instances of p∧q is the intersection of one set of worlds (the content of p) and another set of worlds (the content of q). The condition under which the entailment (p1, p2…pn)⇒q is true is the situation in which the intersection of the possible worlds such that p1, p2…pn is a subset of the possible worlds such that q[47].

On Gibbard's revised account of modal semantics, normative and moral statements differ from nonnormative and nonmoral statements insofar as their content is spelled out by a pair of a possible worlds assignment and a set of norms. Gibbard defines a set of norms as

> the end result of the ways the general normative principles a person accepts combine, weigh against each other, and override each other (Gibbard 1990, 95).

[47] For the sake of simplifying this discussion, I am not spelling out the complete modal semantic account of all of the connectives. Drier (1996) has a complete account in his excellent summary of Gibbard's views.

As mentioned above, the content of a moral utterance consists of a pair of possible worlds and a set of norms. The content of the sentence "Lying is wrong" is the set of pairs of possible worlds and systems of norms that forbid lying.

Gibbard's account attempts to solve the embedding problem by drawing on the resources of the modal semantics account of the logical connectives. We assess the consistency of a person's normative commitments using an adapted version of these resources. Gibbard presents this view in terms of the following formalism: We assess commitments with respect to a pair of a possible world and a set of norms. If a person accepts norms that are in conflict, that person has inconsistent moral views. If a person accepted L1 and L2 but rejected L3, this would result in a person holding an inconsistent set of norms. To spell this out in detail, the acceptance of norms expressed by L1 and L2 compels one to accept L3:

L1: If lying is always wrong, then lying on your resume to get a better job is wrong (w, n1).

L2: Lying is always wrong (w,n2)

L3: Lying on your resume to get a better job is wrong (w, n3)

According to the account of entailment detailed above, the intersection of the norms expressed by L1 and L2 is a subset of the set of norms expressed by L3. The person who accepted L1 and L2 but rejected L3 would thus both accept L1 and L2 and reject L1 and L2. As with Blackburn's new account, this would result in an inconsistency. It is not entirely clear why we should consider the inconsistency captured by Gibbard's account a *logical* inconsistency rather than a practical one—For this reason, Gibbard's account raises a worry quite similar to the one that has been raised against Blackburn's account.

There is also some doubt as to whether Gibbard's account genuinely solves the problem. Blackburn contends that Gibbard's notion of a system of norms assumes far too much, for it takes for granted a notion of logical consistency among norms. How does Gibbard show that norms can be described using the set-theoretical resources deployed by possible world semantics? Blackburn contends that Gibbard provides no explanation of why it would be inconsistent to accept a particular norm N and reject that norm N:

> the problem seems only to be sidestepped if we start off representing attitudes in a structure defined by

> notions of consistency and inconsistency (Blackburn 1993b, 949)

There are also reasons to worry that the possible worlds framework that forms the basis for Gibbard's account is incorrect generally. A significant number of philosophers are skeptical of the idea that propositions are to be spelled out in terms of possible worlds, and only a philosopher inclined to accept such an account of propositions would find Gibbard's account attractive. It would be hard to see, for example, how a philosopher who held the view that propositions consist of Fregean senses could accept Gibbard's overall theory.

It is worth noting here, to relate this discussion to the argument presented in this chapter, that none of the attempted solutions presented above conflict in any way with deflationism or minimalism regarding truth. If there is an embedding problem, the minimalist could embrace any of the approaches taken by Blackburn or Gibbard to the problem. These theories concern the meaning of expressions such as the logical connectives, a separate issue from issue of the proper account of truth.

There has, however, been a dispute as to whether minimalism regarding truth and related minimalist theories of facts and properties bear on the question of how to

solve the embedding problem. Horwich 1993 argues that, if accounts are given of the semantics of moral terms according to which moral terms are ordinary predicates, and the aforementioned minimalist accounts are accepted, then there is a quite simple solution to the embedding problem. A use theory of meaning for moral predicates (of the kind discussed in Chapter 3 of this dissertation) allows us to offer an account according to which moral predicates are ordinary predicates, and hence would lay the groundwork for such a solution.

Dreier 1996 criticizes Horwich's account, misleadingly summarizing Horwich's view as one according to which when minimalism is accepted, "the Embedding Problem vanishes" (Drier 1996, 32). In response to this view, which is not Horwich's own, Dreier contends that the embedding problem would still be with us even if minimalism regarding truth, facts, and properties were accepted. Gibbard endorses Dreier's criticism of Horwich, and implicitly endorses this misleading characterization of Horwich's view:

> My own view of the matter is that (i) there is a[n embedding] problem, (ii) minimalism about truth won't take care of it, and (iii) the solution I have been offering does solve the problem (Gibbard 2003, 63).

Pace Dreier and Gibbard, I will contend that Horwich's solution of the embedding problem is not susceptible to the objections Dreier offers against this account. Thus the doubts raised by Gibbard and Dreier regarding Gibbard's point ii—the claim that minimalism could play some role in the solution of the embedding problem—are not warranted doubts.

The presupposition of philosophers who hold that minimalism can play no role in the solution of the embedding problem is that there is a major difference between the function of moral language and the function of non-moral language. Blackburn puts this point as follows:

> To those impressed by the Frege-Geach [embedding] problem, it is as if there is an abyss between the simple states of mind the expressivist relies upon, and real *judgment* (Blackburn, 1992).

For Blackburn, a real judgment involves a belief suitably related to a property, whereas moral non-judgments involve desires that are related to no properties whatsoever. This purported abyss is a result of a problem that was noted earlier in this dissertation, the tendency to cast expressivism as the denial of the existence of moral facts and properties. If it is granted that there are legitimate, minimalist notions of facts and properties (as

spelled out in Horwich 1993 and Chapter 4 of this dissertation), and moral predicates are indeed genuine predicates, then this is not the correct way to spell out an expressivist view. As I argued in Chapter 4, the best formulation of expressivism is as a view denying the existence of robust moral facts and robust moral properties, but not denying the existence of moral facts and properties. What is required for there to be moral facts and properties in this sense is for there to be true moral utterances, and for moral predicates to be genuine predicates. The account of the semantics of moral terms given in Chapter 3 of this dissertation shows how one can appeal to a use theory of meaning to characterize moral predicates as genuine ones.

If we accept such accounts, it seems Blackburn's Fregean abyss is no longer quite so deep. We can appeal to a use theory of meaning to present an account of the meaning of moral terms, an account that does not require denying that moral predicates are genuine ones. There is no difference between moral and non-moral language insofar as both involve the attribution of properties and assertion of facts. We can appeal to a uniform account of the logical connectives—one cast in terms of a use theory of

meaning—to explain all arguments containing moral and non-moral language.

As noted above, Dreier claims that Horwich's account fails to make the Frege-Geach problem disappear, and offers an example that is supposed to show there remains a problem to be solved. This is not a matter about which Horwich and Dreier are genuinely in disagreement. Horwich's view is not the view that minimalism would make the embedding problem disappear—According to Horwich, there is an embedding problem, and the resources of a use theory of meaning and minimalism show how to solve this problem.

That there is in fact no genuine disagreement between Dreier and Horwich can be illustrated by appeal to the example Dreier uses in his arguments against Horwich. Dreier asks us to imagine a predicate, 'hiyo.' The predicate 'hiyo' is used in a specific construction in order to accost someone. If I say 'Susan is hiyo,' then I am performing the speech act of accosting Susan. This is intended to be parallel to the expressivist account of the meaning of moral terms such as 'right,' which serves a similar purpose in similar contexts: 'Redressing inequalities in the distribution of wealth is right' is, according to the expressivist, used to perform the speech

act of expressing approval of redressing inequalities in the distribution of wealth.

Dreier rightly points out that we would balk at accepting the following argument as valid:

Argument 1

P1: If it is snowing, then Susan is hiyo.

P2: It is snowing.

Therefore, C: Susan is hiyo.

The difficulty is that one cannot have the conclusion of an argument be a speech act of accosting. In order to tell the difference between this illegitimate argument and the legitimate argument below, Dreier claims that we need a solution to the Frege-Geach problems along the lines of Gibbard's:

Argument 2

P1: If redressing inequalities in the distribution of wealth is right, then there should be progressive taxation.

P2: Redressing inequalities in the distrubution of wealth is right.

Therefore, C: There should be progressive taxation.

Gibbard's account is preferable, according to Dreier, given that it blocks argument 1 and allows argument 2. It does so by being cast in terms of normative vocabulary and

norms: Only sentences containing normative language are candidates for Gibbard's analysis. Argument 1 does not fit these criteria, hence Gibbard's account correctly rules it out.

Should someone convinced of Horwich's account be moved by such an argument? It is not clear that this would be so. It is crucial to keep in mind that Horwich's position is not the strawman attributed to him by Dreier—It is not the view that minimalism makes the embedding problem disappear. Rather, Horwich claims that, in light of an account according to which moral predicates are genuine predicates, and in light of minimalism and a conceptual role account of logical connectives, we can assimilate moral and non-moral arguments. What Horwich and the minimalist need not accept, as Drier claims that they must, that 'hiyo' could be stipulated to be a genuine predicate. Thus the minimalist can agree with Dreier that any attempt to stipulate the meaning of 'hiyo' in the way described in Dreier's example would fail. Thus this example does not present a reason to reject Horwich's solution of the Frege-Geach problem.

It may be thought that Horwich's approach to the Frege-Geach problem involves rejecting the theses most distinctive of contemporary expressivism, the detailed,

complex, and interesting accounts of the semantics of the logical connectives offered by Blackburn and Gibbard. Perhaps this would be another reason to think the minimalist account is incompatible with expressivism[48]. I do not think this is the case. First, this resolution of the problem is not in conflict with any of the essential elements of expressivism (as detailed in the beginning of this chapter, a characterization drawn from the works of these contemporary expressivists). Also, as the Frege-Geach problem was taken, before the presentation of attempted substantive solutions, to be a problem for expressivism, this approach to the problem would seem to pave the way for expressivism rather than to destroy the theory from within.

[48] I owe this objection to Jonathan Adler.

CHAPTER 7

CONSTRUCTIVIST ETHICS AND DEFLATIONARY TRUTH

The expressivist theory discussed in the previous chapter has, as I noted above, a historical lineage dating back to the work of the great philosopher David Hume. Among the major meta-ethical alternatives to expressivism are constructivist theories, theories that have an origin in the work of another great philosopher, Immanuel Kant. Constructivist theories have been developed by a number of philosophers (Korsgaard 1996, Ross 1991, Ross 1998, Scanlon 1998). The best-known and most influentual recent formulation of such a view is due to John Rawls (1980). There has been a great deal of discussion of the relationship between expressivism and theories of truth. In stark contrast, there has been very little discussion[49] of the relationship between constructivist theories and theories of truth. This is especially surprising in light of the fact that constructivism presents a *prima facie* problem quite similar to the incompatibilist worry discussed in the previous chapter.

[49] An exception that will be discussed later in the chapter is the discussion of constructivism and (pragmatist) truth in Cheryl Misak's *Truth, Politics, Morality*.

The *prima facie* problem arises from Rawls's claim in his essay "Kantian Constructivism in Moral Theory" that principles of justice ought not to be considered true or false, but rather reasonable or unreasonable. Rawls claims:

> Given the various contrasts between Kantian constructivism and rational intuitionism, it seems better to say that in constructivism first principles are reasonable or unreasonable than that they are true or false… (Rawls 1980, 355).

There is a conflict between this claim and the account of the truth of moral principles that ought to be accepted by the deflationist. In light of the arguments presented in Chapters 1 and 2 of this dissertation, a deflationist ought to consider the "first principles" discussed by Rawls to be capable of being straightforwardly true or false. Thus it seems that there is a conflict between constructivism and deflationism: Deflationist theories such as minimalism give us reason to attribute truth to principles of justice, and thus are in tension with Rawls's Kantian constructivist theory, which withholds the attribution of truth from these principles.

It is important to note that the apparent conflict between deflationism and constructivism is not the only

relevant issue here. There are also general questions regarding the notion of truth employed in Rawls's explanation of constructivism: How plausible is the account of truth attributed by Rawls to the rational intuitionist? Is this an account of truth that any theorist of truth, whether deflationist or inflationist, would accept? While Rawls has stated that he considers moral and political theory to be autonomous, in important respects independent of metaphysics[50], are the claims made by Rawls regarding the truth of moral principles required for maintaining this autonomy?

In the first section of this chapter, **The Essence of Constructivism**, I will describe constructivist theories, indicating what is distinctive of such theories. This section will draw on the account of the relationship between truth, metaphysics, and morality developed in earlier chapters of this dissertation. In the following section, **Rawls's Kantian Constructivism**, I will detail Rawls's particular formulation of the Kantian constructivist theory. In spelling out Rawls's Kantian constructivism, I will present Rawls's reasons for claiming

[50] Rawls argues for the autonomy of political philosophy in "Justice as Fairness: Political not Metaphysical" (Rawls 1985).

that truth ought not to be attributed to principles of justice.

In the section entitled **Resolving the *Prima Facie* Conflict**, I argue that the Rawlsian argument against attributing truth to principles of justice only succeeds if one holds an account of truth (generally) and moral truth (specifically) that would be rejected by deflationists, among others. Thus a deflationist regarding truth who holds a constructivist meta-ethical view need not hold, as Rawls does, that truth ought not to be attributed to principles of justice. In the final section of the chapter, **What a Constructivist Ought to Say About Truth**, I argue that a position combining a constructivist meta-ethics with a deflationary account of truth such as minimalism has great advantages over a view that does not attribute truth to moral principles, reiterating arguments presented in previous chapters.

The Essence of Constructivism

On a constructivist theory, the correctness of moral judgments is not to be explained in terms of correspondence of these judgments to a realm of robust moral properties or facts. In this respect, constructivism is similar to expressivism. These views share a lack of appeal to robust moral facts and robust moral properties. The expressivist,

as noted above, denies the existence of robust facts and properties. The constructivist, such as Rawls, neither asserts nor denies the existence of robust moral facts and properties. Unlike the expressivist, who is in a way an atheist about robust moral facts and properties, the constructivist is agnostic. Rawls expresses this agnosticism when he writes regarding the "controversy between realism and subjectivism[51] about the status of moral values" that the constructivist "neither asserts nor denies these doctrines" (Rawls 1985, 395).

Another central respect in which constructivist theories differ from expressivist theories (generally) is in terms of the account given of the derivation of proper moral claims. On a constructivist view, proper moral claims are the outcome of a certain procedure, typically a hypothetical decision procedure. In order to clarify the nature of constructivism, I will provide two examples of such a procedure.

On Kant's original formulation of constructivism in the *Groundwork for the Metaphysics of Morals*, proper moral claims are the outcome of the categorical imperative procedure. As is well known, this procedure is spelled out

[51] I am assuming here that Rawls is using the term "subjectivism" to refer to expressivism and similar views here.

in several, perhaps equivalent, ways. To simplify the discussion, I will discuss only one of these formulations of the categorical imperative. On the universal law formulation of the categorical imperative procedure[52], an individual begins by considering the maxims on which she acts, the reasons she has for doing what she plans to do. She then asks herself whether she could will the maxim on which she acts into a universal law.

She may ask herself, for example, whether she should lie on her resume in order to get a better job. If she cannot will this maxim into a universal law, then the maxim fails the procedure and is not an acceptable moral principle. Those maxims that do not result in a conflict when willed into universal laws are proper moral principles. Kant, as is well known, would contend that a person could not will lying on one's resume to get a better job into a universal law. What makes this Kantian account a constructivist view is the account of how moral principles are derived: The principles are the outcome of the procedure, and the outcome of the procedure depends upon the reasoning of the individual contemplating the

[52] I am selecting this version of the categorical imperative as the version to be discussed in the text for the sake of simplicity. This is not to suggest that this formulation is the most interesting or important formulation of the CI procedure.

maxim. In this way, moral principles are discovered by those persons who apply the categorical imperative procedure to their reasons for acting.

A different procedure for testing principles is discussed by Rawls in his *A Theory of Justice*. It is very important to note that Rawls's account is not intended to be an account of moral principles generally (as Kant intends his account) but only an account of political principles of justice, principles to which one would appeal when determining the structure of a society[53]. On Rawls's account, the hypothetical procedure for determining principles of justice is one in which a group of people, a society, discusses such principles from behind a veil of ignorance. The veil of ignorance is a restriction on the information a person has about herself and her situation: She does not know factors such as her social class, race, gender, sexual orientation, religion, or her own moral views. Without this information, these people decide on basic principle of justice for the society. Rawls contends that behind this veil of ignorance, the discussion of

[53] Philosophers influenced by Rawls have spelled out constructivist views that are not restricted to political principles. T.M. Scanlon's book *What We Owe to Each Other* presents a constructivist account of ethics similar in many respects to the constructivist account of politics offered by Rawls.

principles of justice would yield the following two principles:

> First: each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others.
>
> Second: Social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage and (b) attached to positions and offices open to all (Rawls 1971, 53)

These principles have been formulated in a number of different ways, in *A Theory of Justice* as well as other works such as *Political Liberalism*. For the sake of simplicity here, I will discuss Rawls's view with respect to this early, relatively simple formulation of the two principles throughout this chapter.

As with Kant's view, the crucial element that makes this account constructivist is that the correct moral principles are what would be the result of a certain hypothetical procedure carried out by people. For Rawls, the procedure is not to be thought of as one that would be applied by an individual to her own moral decision making (unlike Kant), but rather as a procedure hypothetically carried out by the society as a whole. The political

principles eventually produced by the procedure are the result of a community of human beings reasoning with each other regarding the way a society should be structured. This social aspect of the Ralwsian procedure is crucial due to Rawls's aim to give an account of political justice. His aim is to give an account of the principles members of a society could justify to each other. Invoking a social procedure is a way to prevent one group in a society from imposing a political structure on another group that would not accept such a system.

Moral and political constructivists appeal to such procedures in order to explain the objectivity of moral and political judgments. This account is not a semantic account, an account of the meaning of moral terms. Rather, these accounts aim to show how moral and political judgments are justified. Such philosophers contend that a Humean or expressivist account, which does not generally involve appeal to the constructivist procedures, cannot explain why it is that certain principles have a justification that should convince anyone doubtful of the truth of such principles. How can the expressivist explain why slavery would be wrong even if I and everyone else did

not feel that way?[54] That the principles offered by the constructivist are those that would be the outcome of the specified hypothetical decision procedure is key in the explanation of why such principles are justified. On the Rawlsian account, objective justification of political principles is rooted in more basic facts regarding human rationality and the role played by luck in determining features of a person such as race, gender, and social class.

In response to this constructivist claim to objectivity, expressivists such as Simon Blackburn have contended that expressivism is equally capable of if not better suited for explaining objectivity than the constructivist account (Blackburn 2001). Blackburn contends that a constructivist view tends toward

---

[54] This criticism of expressivism is due to Thomas Nagel. I am grateful to Nagel for discussion of these matters as well as the suggestion, in class, that non-correspondence accounts of truth are best suited for accounting for the truth of moral statements on a constructivist meta-ethical theory. While Nagel has not discussed the issue of the relation between theories of truth and issues in metaethics in detail in his published work, his remarks on moral truth in *The View from Nowhere* suggest that he favors an epistemic account of moral truth as a replacement for a correspondence theory: "I do not believe that the truth about how we should live could extend radically beyond any capacity we might have to discover it…" (Nagel 1986, 139). For the reasons I have detailed in this dissertation, I think a deflationary theory of truth, rather than an epistemic account of truth, is best suited to account for truth (generally) and moral truth (specifically).

relativism. I do not think that this criticism is correct. It is a key aspect of Rawls's view that the two principles of justice would be the outcome of the constructivist procedure whenever such a procedure is carried out by a group of rational individuals. There is no suggestion in Rawls that the principles of justice are only true to one society, but not true to another. Perhaps Blackburn is concerned about Rawls's refusal to attribute truth to the principles of justice, a matter to which I return in the following section.

Rawls's Kantian Constructivism

As noted above, Rawls contends that a constructivst account ought not to claim that the principles that are the result of the procedure of construction are true or false. In order to make entirely clear what account of the truth of moral statements Rawls is rejecting, it is necessary to focus on his description of rational intuitionism in "Kantian Constructivism in Moral Theory."[55] Rawls

[55] I am grateful to Jonathan Adler for pointing out the importance of Rawls's rejection of rational intuitionism for understanding the discussion of the contrast between true and reasonable statements. It ought to be noted that the distinction Rawls makes between true statements and reasonable statements is not the same as the distinction between the reasonable and the rational. As is noted in the revised version of "Kantian Constructivism in Moral Theory" in Rawls's *Collected Papers*, in the discussion of truth of moral principles, "Here 'reasonable' is used, as

characterizes intuitionism, a view he attributes to Plato, Aristotle, Sidgwick, Moore, and W.D. Ross (among others). This is a lengthy characterization, but it is one that is important to discuss in detail in order to fully understand Rawls's discussion of truth:

> For our purposes here, rational intuitionism may be summed up by two theses: first, the basic moral concepts of the right and the good, and the moral worth of persons, are not analyzable in terms of nonmoral concepts (although possibly analyzable in terms of one another); and, second, first principles of morals (whether one or many), when correctly stated, are self-evident propositions about what kinds of considerations are good grounds for applying one of the three basic moral concepts, that is, for asserting that something is (intrinsically) good, or that a certain action is the right thing to do, or that a certain trait of character has moral worth. These two theses imply that the agreement in judgment which is so essential for an effective public conception of justice is founded on the recognition of self-evident

---

explained later…in contrast with 'true' as understood in rational intuitionism, and not, as previously…with 'rational,' as in the notion of rational autonomy" (Rawls 1980, 340-41).

> truths about good reasons. And what these reasons are is fixed by a moral order that is prior to and independent of our conception of the person and the social role of morality. This order is given by the nature of things and is known, not by sense, but by rational intuition. It is with this idea of moral truth that the idea of first principles as reasonable will be contrasted (Rawls 1980, 343-344).

Thus it is clear that Rawls has a particular, detailed account of moral truth in mind here. It is helpful to separate this account of moral truth into its component elements.

The rational intuitionist account of truth, which I will call RIT throughout the rest of this chapter, makes claims regarding the semantics of moral terms, the metaphysics of morality, and the epistemology of morality. One component of RIT, the semantic component, is a theory regarding the definition of moral terms and concepts. Moral terms and moral concepts are unanalyzable, or capable only of a circular analysis. No definition can be given of a basic moral term such as 'good,' 'just,' or 'right.'

The metaphysical component of RIT is the claim that there is an independent realm of robust moral facts and properties. There is, on the RIT account of moral

metaphysics, a "moral order" that is "prior to and independent of our conception of the person and the social role of morality…given by the nature of things" (ibid.). In contrast to this RIT account of the metaphysics of moral entities, there is no need in constructivism to appeal to such a realm of moral entities in the justification of moral statements.

The epistemic component of RIT involves how moral claims are known: these claims are known by a process of intuition. It is quite obscure what rational intuition is supposed to be, and I will not try to spell it out in detail in this chapter. Let it suffice to say that the RIT epistemic component requires that moral claims are self-evident, claims that require no further justification.

Resolving the *Prima Facie* Conflict

Having spelled out the RIT account of moral truth rejected by Rawls, it is clear that the conflict noted above between deflationism and constructivism is only an apparent one. Appealing to the descriptions of deflationary theories given in Chapter 1, it is fairly obvious that none of these deflationary theories of truth contain commitments regarding the semantics, metaphysics, or epistemology of morality anything like the commitments required for the attribution of truth on the RIT account.

These deflationary accounts of truth are not intended to settle the matters involved in the commitments of RIT.

Regarding the semantic component of RIT, deflationist views might involve an account of the meaning of 'true,' and these theories cohere best (as discussed in Chapter 3) with the framework of a use theory of meaning, but no deflationary theory by itself settles the matter of whether 'good' and 'right' are analyzable or unanalyzable. The epistemic component of RIT is equally independent of a deflationary account of truth. Whether moral principles are self-evident, or whether their justification is explained in terms of a constructivist procedure is not accounted for by a deflationary theory of truth. Deflationary theories neither commit one to a particular account of moral justification, nor do they exclude any account. From the discussion of metaphysics in earlier chapters of this dissertation, it ought also to be clear that the matter of the truth of moral statements is a distinct matter from the metaphysical component of RIT, the commitment of the rational intuitionist to robust moral facts and properties.

If Rawls had taken account of this independence of the commitments of deflationist accounts of truth from the commitments of RIT in his discussion of the truth of moral

and political principles, then he could quite easily have claimed that, for a deflationist regarding truth, principles of justice are true as well as reasonable. While Rawls made the distinction between reasonable statements and true statements as a way to characterize the distinction between his view and rational intuitionism, an account such as the one in this chapter setting out each of these issues (truth, semantics, metaphysics, epistemology) separately makes it quite clear in what respects Rawls's view differs from that of the rational intuitionist.

Perhaps Rawls thought, when he wrote "Kantian Constructivism in Moral Theory," that the only way to characterize moral statements as true would be in terms of a correspondence theory. If this were the case, then it would be necessary for one to appeal to the existence of robust moral facts and properties in order to properly attribute truth to moral utterances. In light of the more recent development of deflationist and minimalist theories of truth, we can see that this apparent background assumption made by Rawls is unwarranted—there is a deflationary alternative to the correspondence theory that would allow us to attribute truth to moral utterances.

In order to see that there is no conflict between Rawls's view and deflationism, we need only notice that it

is not necessary to mention truth to make the distinction between Rawls and the rational intuitionist. All that is necessary to distinguish between Rawls and the rational intuitionist is to note that Rawls does not take a stand on the semantic component of RIT, is agnostic regarding the metaphysical component of RIT, and rejects the epistemological component of RIT. These difference between Rawlsian constructivism and rational intuitionism thus can be stated without any mention of the notion of truth.

That Rawls need not have made the claim that moral principles are incapable of being true in order to spell out a distinctive constructivist view has been noted previously by pragmatist philosophers. In her book *Truth, Politics, Morality*, Cheryl Misak advocates a pragmatist account of truth, an account derived from the work of Charles Sanders Peirce. According to Misak, true statements are those statements that would be the result of an inquiry that could not be improved upon. While there are reasons to be doubtful of the correctness of such an account of truth, as was noted in Chapter 1[56], the

[56] There are two reasons to be dubious of Misak's account. First, it does not seem intuitively correct that claims that could not possibly be settled by human inquiry are neither true nor false. It may perhaps, for example, be impossible for humans ever to fully understand or discover the nature of dark matter. Does this really mean that

pragmatist account shares with deflationary accounts one significant feature, namely that both accounts reject the idea that truth is a "metaphysically weighty" notion. Unlike a correspondence theory, no requirement is made by the pragmatist or the deflationist that some kind of entity, fact, or truthmaker must exist in order to make each true sentence true. Thus both pragmatism and deflationism are fully consistent with and independent of Rawls's agnosticism regarding the existence of moral facts and properties.

Misak suggests that Rawls is thinking of a metaphysically weighty account of truth of the kind rejected by pragmatists (and deflationists) when he claims that moral principles are reasonable or unreasonable, but neither true nor false. According to Misak:

> [Rawls] does advise us to drop truth talk from politics and from political philosophy. Political liberalism is unmetaphysical in that it does not claim any kind of truth for itself. A political conception of justice 'does without the concept of truth' (Rawls

---

there are no truths or falsehoods about the aspects of dark matter we cannot discover? Second, as a result of the first point, pragmatism requires a revision of the principle of excluded middle, a radical result in logic not required by a deflationist view. A deflationist account shares many of the advantages of pragmatism without requiring such a revision.

> 1993, 94)…If we take the pragmatist's advice and refuse to think of truth as a metaphysical, all-or-nothing concept, perhaps Rawls would be less wary of thinking that political liberalism is true. Political liberalism, it might be argued, is the best arrangement, the best way to conduct our public lives. If that is right then, on the pragmatist's view of truth, political liberalism is true (Misak 2000, 21).

When it is also noted that a deflationist account of truth does not involve the metaphysical commitments Rawls associates with truth, then it should be clear that the essence of a Kantian constructivist view does not require withholding the attribution of truth from the principles of justice. The matters Rawls is most concerned with here are the autonomy of moral and political theory from metaphysics, and the proper account of the justification of moral principles, and deflationist and pragmatist theories demonstrate that truth, metaphysics, and moral justification are at least conceptually distinct matters.

What a Constructivist Ought to Say About Moral Truth

In this final section of the chapter and of the dissertation, I will reiterate the reasons presented in Chapter 2 for attributing truth to moral utterances. That any deflationist account of truth allows for the

attribution of truth to moral utterances is an attractive feature of such accounts of truth, and it presents a reason for moral constructivists to accept a deflationist account.

The reasons for attributing truth to moral utterances are as follows. Attributions of truth to moral utterances are to be expected given that moral discourse is at least on the surface quite similar to other types of discourse in which truth is commonly attributed to utterances. The attribution of truth to moral utterances is a feature of ordinary practice; People are quite willing to make assertions such as 'It is true that rape is wrong,' or 'It is false that wanton killing of innocents is permissible.' The role of arguments in moral reasoning also cannot be accounted for in a proper and simple fashion without the attribution of truth to moral utterances. As I have noted several times in this dissertation, given that validity is defined in terms of truth, rejection of the attribution of truth to moral utterances makes it difficult or impossible to explain how moral utterances could be valid or invalid. A final reason for attributing truth to moral utterances is that it would complicate theories of truth unnecessarily if we had to revise them in order to deny that truth ought to be attributed to moral utterances.

In light of these considerations, the best formulation of a constructivist view is the one that rejects Rawls's claim that moral principles are neither true nor false. As long as it is made clear that the deflationist account of truth does not carry with it the semantic, metaphysical, and epistemic commitments of the rational intuitionist account of truth (RIT), rejection of this aspect of Rawls's formulation of constructivism does not amount to a rejection or even a conflict with the essence of Rawls's view.

WORKS CITED

Adams, Fred and Ken Aizawa. 1994. "Fodorian Semantics." In Stich, Stephen P. and Ted A. Warfield 1994.

Austin, J.L. 1950. "Truth" in Austin 1970: 117-33.

___. 1970. *Philosophical Papers*. Second Edition. Edited by J.O. Urmson and G.J. Warnock. London: Oxford University Press.

Ayer, A.J. 1936. *Language, Truth, and Logic*. London: Dover.

Blackburn, Simon. 1984. *Spreading the Word*. Oxford: Clarendon Press.

___. 1988a. "Attitudes and Contents." *Ethics* 98: 501-17.

___. 1988b. "How to be an Ethical Antirealist" in Darwall, Gibbard, and Railton 1997: 167-178.

___. 1993a. *Essays in Quasi-Realism.* New York: Oxford University Press.

___. 1993b. "Gibbard on Normative Logic" in Enrique Villanueva, ed. *Philosophical Issues 4: Naturalism and Normativity.* Atascadero, Calif.: Ridgeview.

___. 1993c. "Realism: Quasi or Queasy" in Haldane and Wright, eds. *Reality.* Oxford: Oxford University Press.

___. 1998. "Wittgenstein, Wright, Rorty, and Minimalism." *Mind* 107: 157-82.

___. 1999. "Is Objectively Moral Justification Possible on a Quasi-Realist Foundation?" *Inquiry* 42: 213-28.

___. Manuscript. "The Semantics of Non-Factualism, Non Cognitivism, Quasi-Realism."

Blanshard, Brand. 1939. *The Nature of Thought*. Volume 2. London: Allen and Unwin.

Block, Ned. 1986. "Advertisement for a Semantics for Psychology" in P. French, T. Uehling, and H. Wettstein eds., *Midwest Studies in Philosophy* 10. Minneapolis, Minn.: University of Minnesota Press.

Boghossian, Paul. 1990a. "The Status of Content." *Philosophical Review* 99: 157-84.

___. 1990b. "The Status of Content Revisted." *Pacific Philosophical Quarterly* 71:264-78.

Boghossian, Paul and Christopher Peacocke. 2000. *New Essays on the A Priori*. Oxford: Clarendon Press.

Boyd, Richard. 1988. "How to be a Moral Realist" in Darwall, Gibbard, and Railton 1997.

Brandom, Robert. 1988. "Pragmatism, Phenomenalism, and Truth Talk" in P. French, T. Uehling, and H. Wettstein eds., *Midwest Studies in Philosophy* 12. Minneapolis, Minn.: University of Minnesota Press

___. 1994. *Making it Explicit*. Cambridge, Mass.: Harvard University Press.

Brentano, Franz. 1924-25. *Psychologie vom empirischen Standpunkt.* Edited by Oskar Kraus. 2 vols. Leipzig: F. Meiner.

Brighouse, M. 1990. "Blackburn's Projectivism—An Objection." *Philosophical Studies* 59: 225-33.

Campbell, Keith. 1981. "The Metaphysic of Abstract Particulars." *Midwest Studies in Philosophy* 6: 477-488.

Chang, Ruth. 2002. "The Possibility of Parity." *Ethics* 112: 659-688.

Chrisman, Matthew. "*Thinking How to Live*." *Ethics* 115: 406-411.

Church, Alonzo. 1956. *Introduction to Mathematical Logic*. Vol. 1. Princeton: Princeton University Press.

Copp, David. 2001. "Realist-Expressivism: A Neglected Option for Moral Realism." *Social Policy and Philosophy* 18: 1-43.

Craig, Edward, ed. 1998. *The Routledge Encyclopedia of Philosophy*. London and New York: Routledge.

Cullity, Garrett. 1998. "Moral Judgement." In Craig, Edward, ed. 1998, vol. 6: pp. 511-514.

Darwall, Stephen, Allan Gibbard, and Peter Railton. 1997. *Moral Discourse and Practice: Some Philosophical Approaches*. New York: Oxford University Press.

___. 1992. "Toward *Fin de siècle* Ethics: Some Trends" in Darwall, Gibbard, and Railton 1997: 3-47.

Davidson, Donald. 1984. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press

Devitt, Michael. 1980. "'Ostrich Nominalism' or 'Mirage Realism'?" in Mellor and Oliver 1997: 93-100.

___. 1996. "The Metaphysics of Nonfactualism" in J. Tomberlin, ed. *Philosophical Perspectives* 10.

___. 1997. *Realism and Truth*. Second Edition. Princeton: Princeton University Press.

___. 2001. "The Metaphysics of Truth" in Lynch 2001: 579-611.

___. 2002. "Moral Realism." *Croatian Journal of Philosophy* 2: 1-15.

Dewey, John. 1929. *The Quest for Certainty*. New York: G.P. Putnam's Sons.

Divers, J. and Miller, A. 1994. "Why Expressivists about Value Should Not Love Minimalism about Truth." *Analysis* 54: 12-19.

Dreier, James. 1996. "Expressivist Embeddings and Minimalist Truth." *Philosophical Studies* 83: 29-51.

___. 1999. "Transforming Expressivism." *Noûs* 33: 558-72.

___. 2004. “Meta-ethics and the Problem of Creeping Minimalism.” *Philosophical Perspectives* 18: 23-44.

Dretske, Fred. 1988. *Explaining Behavior*. Cambridge, Mass.: MIT Press.

Dummett, Michael. 1959. “Truth” in *Truth and Other Enigmas*. London: Duckworth.

___. 1978. *Truth and Other Enigmas*. Cambridge, Mass.: Harvard University Press.

___. 1982. “Realism.” *Synthese* 52: 55-112.

Dworkin, Ronald. 1996. “Objectivity and Truth: You’d Better Believe It.” *Philosophy and Public Affairs* 25: 87-139.

Field, Hartry. 1972. “Tarski’s Theory of Truth.” In Field 2001: 3-29.

___. 1977. “Logic, Meaning, and Conceptual Role.” *Journal of Philosophy* 69: 379-409.

___. 1986. “Correspondence Truth, Disquotational Truth, and Deflationism.” In Lynch 2001: 483-503.

___. 1994. “Disquotational Truth and Factually Defective Discourse.” *Philosophical Review* 103: 405-52.

___. 2001. *Truth and the Absence of Fact*. Oxford: Clarendon.

Fine, Kit. 2001. “The Question of Realism.” *Philosopher’s Imprint* 1: 1-30.

Fodor, Jerry. 1990. *A Theory of Content and Other Essays*. Cambridge, Mass.: MIT Press.

Fodor, Jerry, and Ernest Lepore. 1991. *Holism: A Shopper’s Guide*. Oxford: Blackwell.

___. 2001. “Brandom’s Burdens.” *Philosophy and Phenomenological Research* 63: pp. 465-482.

Frege, Gottlob. 1918-19. "Thought." In Frege, Gottlob 1997: 325-45.

___. 1997. *The Frege Reader*. Edited by Michael Beaney. Oxford: Blackwell Publishers.

Geach, Peter. 1964. "Assertion." *Philosophical Review* 74: 449-465.

Gibbard, Allan. 1986. "An Expressivistic Theory of Normative Discourse." *Ethics* 96: 472-85.

___. 1990. *Wise Choices, Apt Feelings*. Oxford: Oxford University Press.

___. 1991. "Constructing Justice." *Philosophy and Public Affairs* 20: 264-79.

___. 1992a. "Moral Concepts: Substance and Sentiment" in Tomberlin, James E. ed., *Philosophical Perspectives 6: Ethics*. Atascadero, Calif.: Ridgeview.

___. 1992b. "Reply to Blackburn, Carson, Hill, Railton." *Philosophy and Phenomenological Research* 52: 969-980.

___. 1993. "Reply to Blackburn." In Enrique Villanueva, ed. *Philosophical Issues 4: Naturalism and Normativity*. Atascadero, Calif.: Ridgeview.

___. 2003a. "Reasons Thin and Thick." *The Journal of Philosophy* 100: 288-304.

___. 2003b. *Thinking How to Live.* Cambridge, Mass.: Harvard University Press.

Grover, Dorothy. 1992. *A Prosentential Theory of Truth*. Princeton: Princeton University Press.

Grover, Dorothy, Joseph Camp, and Nuel Belnap. 1975. "A Prosentential Theory of Truth." In Grover 1992: 70-120.

Hale, Bob. 1986. "The Compleat Projectivist." *The Philosophical Quarterly* 36: 65-84.

___. 1993. "Can There Be a Logic of Attitudes?" in Haldane and Wright, eds. *Reality.* Oxford: Oxford University Press.

Hare, R.M. 1997. *Sorting Out Ethics*. Oxford: Clarendon Press.

Harman, Gilbert. 1982. "Conceptual Role Semantics." *Notre Dame Journal of Formal Logic* 23: 242-56.

___. 1987. "(Nonsolipsistic) Conceptual Role Semantics" In Lepore 1987.

Holton, Richard. 2000. "Minimalism and Truth Value Gaps." *Philosophical Studies* 97: 137-68.

Horwich, Paul. 1993. "Gibbard's Theory of Norms." *Philosophy and Public Affairs* 22: 67-78.

___. 1994. "The Essence of Expressivism." *Analysis* 45: 19-20.

___. 1996a. "Realism and Truth" in J. Tomberlin, ed. *Philosophical Perspectives* 10.

___. 1996b. "Realism Minus Truth." *Philosophy and Phenomenological Research* 56: 877-881.

___. 1997. "Realism and Truth." *Ponzani Studies in the Philosophy of the Sciences and the Humanities* 55: 29-39.

___. 1998a. *Meaning*. Oxford: Oxford University Press.

___. 1998b. *Truth*. Oxford: Oxford University Press.

Hume, David. 2000. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.

Hurley, S.L. 1988. *Natural Reasons*. Oxford: Oxford University Press.

Jackson, Frank, Graham Oppy, and Michael Smith. 1994. "Minimalism and Truth Aptness." *Mind* 103: 287-302.

James, William. 1907. "Pragmatism's Conception of Truth." In Lynch 2001: 211-228.

Kant, Immanuel. 1996. *Practical Philosophy*. Translated and edited by Mary J. Gregor. Cambridge: Cambridge University Press.

Katz, Jerrold J. 2001. *Sense, Reference, and Philosophy*. New York: Oxford University Press.

Kölbel, M. 2002. *Truth without Objectivity.* London: Routledge.

Korsgaard, Christine M. 1996. *The Sources of Normativity.* Cambridge, UK: Cambridge University Press.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.

Leeds, Stephen. 1978. "Theories of Reference and Truth." *Erkenntnis* 13: 111-129.

Lepore, Ernest. 1987. *New Directions in Semantics*. London: Academic Press.

Loewer, Barry. 1987. "From Information to Intentionality." *Synthese* 70: pp. 287-317.

Lynch, Michael P., ed. 2001. *The Nature of Truth*. Cambridge, Mass.: MIT Press.

Mackie, John. 1977. *Ethics: Inventing Right and Wrong.* New York: Penguin.

Misak, Cheryl. 2000. *Truth, Politics, Morality.* London: Routledge.

Moore, G.E. 1903. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press.

Nagel, Thomas. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.

___. 1986. *The View from Nowhere*. Oxford: Oxford University Press.

Neale, Stephen. 2001. *Facing Facts*. Oxford: Clarendon Press.

O'Leary-Hawthorne, John, and Huw Price. 1996. "How to Stand up for Non-Cognitivists." *The Australasian Journal of Philosophy* 74: 275-92.

Peacocke, Christopher. 1992. *A Study of Concepts*. Cambridge, Mass: MIT Press.

Peirce, Charles Sanders. 1878. "How to Make our Ideas Clear." In Lynch 2001: 193-209.

Price, Huw. 1994. "Semantic Minimalism and the Frege Point" in S.L. Tsohatzidis, ed. *Foundations of Speech Act Theory*. London and New York: Routledge.

Putnam, Hilary. 1981. *Reason, Truth, and History*. Cambridge: Cambridge University Press.

Quine, W.V. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.

___. 1970. *Philosophy of Logic*. Cambridge, Mass.: Harvard University Press.

___. 1992. *Pursuit of Truth*. Revised Edition. Cambridge, Mass.: Harvard University Press.

Railton, Peter. 1986. "Moral Realism" in Gibbard, Darwall, and Railton 1997.

___. 1998. "Analytic Ethics." In Craig, Edward, ed. 1998, vol. 1: pp. 220-223.

Ramsey, Frank Plumpton. 1990. "The Nature of Truth." In Lynch 2001: 433-446.

Rawls, John. 1980. "Kantian Constructivism in Moral Theory" in Rawls 1999.

___. 1985. "Justice as Fairness: Political not Metaphysical" in Rawls 1999.

___. 1993. *Political Liberalism*. New York: Columbia University Press.

___. 1997. "The Idea of Public Reason Revisited" in Rawls 1999.

___. 1999. *Collected Papers*. Cambridge, Mass.: Harvard University Press.

Ross, Steven. 1991. "The Nature of Moral Facts." *The Philosophical Forum* 22: 243-69.

___. 1998. "The Nature and Limits of Moral Objectivism." *The Philosophical Forum* 24: 28-49.

Russell, Bertrand. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press.

Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.

Searle, John. 1962. "Meaning and Speech Acts." *Philosophical Review* 71: 423-32.

Shafer-Landau, Russ. 2003. *Moral Realism*. Oxford: Oxford University Press.

Schueler, G.F. 1988. "Modus Ponens and Moral Realism." *Ethics* 98: 492-500.

Schiffer, Stephen. 2003. *The Things We Mean*. Oxford: Clarendon Press.

Sellars, Wilfrid. 1954. "Some Reflections on Language Games" in Sellars 1963: pp. 321-358.

___. 1963. *Science, Perception, and Reality*. Atascadero, California: Ridgewood.

Sidgwick, Henry. 1981. *The Methods of Ethics*. Indianapolis: Hackett.

Smith, Michael. 1994a. "Minimalism, Truth Aptitude, and Belief." *Analysis* 54:21-26.

___. 1994b. *The Moral Problem*. Oxford: Blackwell.

___. 1994c. "Why Expressivists about Value Should Love Minimalism about Truth." *Analysis* 54, 1-12.

Soames, Scott. 1984. "What is a Theory of Truth?" in Lynch 2001: pp. 397-418.

Stevenson, C. 1937. "The Emotive Theory of Ethical Terms." *Mind* 46: 14-31.

___. 1963. *Facts and Values*. New Haven: Yale University Press.

Stich, Stephen P. and Ted Warfield. 1994. *Mental Representation: A Reader*. Oxford: Blackwell.

Strawson, P.F. 1950. "Truth." *Proceedings of the Aristotelian Society*, supp. vol. 25: 129-156.

Tarski, Alfred. 1944. "The Semantic Conception of Truth and the Foundations of Semantics." In Lynch 2001: 331-363.

___. 1935. "Der Wahrheitsbegriff in den Formalisierten Sprachen." *Studia philosophica* 1: pp. 261-405.

Walker, Ralph C.S. 1989. *The Coherence Theory of Truth: Realism, Anti-realism, Idealism*. London: Routledge.

Wedgwood, Ralph. 1997. "Non-Cognitivism, Truth, and Logic." *Philosophical Studies* 86: 73-91.

___. 2001. "Conceptual Role Semantics for Moral Terms." *The Philosophical Review* 110: pp. 1-30.

Wittgenstein, Ludwig. 1922. *Tractatus Logico-philosophicus*. London: Routledge and Kegan Paul.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. New York: Macmillan.

Wright, Crispin. 1985. "*Spreading the Word.*" *Mind* 94: 310-19.

___. 1987. "Realism, Anti-Realism, Irrealism, Quasi-realism" in P. French, T. Uehling, and H. Wettstein eds., *Midwest Studies in Philosophy* 12. Minneapolis, Minn.: University of Minnesota Press.

___. 1992. *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press.

___. 1996a. "Precis of *Truth and Objectivity*." *Philosophy and Phenomenological Research* 56: 863-868.

___. 1996b. "Response to Commentators." *Philosophy and Phenomenological Research* 56: 911-941.

Zangwill, N. "Moral Modus Ponens." *Ratio* 5: 177-93.