

OBJECTIVIST COMPATIBILIST UTILITARIANISM

by

NICHOLAS CAMPBELL SLEIGH

B.A., The University of British Columbia

M.A., Simon Fraser University

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Philosophy)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 1993

© Nicholas Campbell Sleigh, 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Philosophy

The University of British Columbia
Vancouver, Canada

Date March 9, 1994

ABSTRACT

In this thesis a version of the ethical theory of utilitarianism is defended. The version defended is called 'objectivist compatibilist utilitarianism', or 'OCU'. On this version, utilitarian metaethics includes the propositions that there is an objective, intrinsic property of goodness entifying, motivating, and grounding ethics, and that act and rule utilitarianism are compatible since under plausible interpretations both true. While this metaethical theory has perhaps not been stated explicitly before, theories in this vein have been popular since the mid-19th century, and have been expounded by philosophers such as J.S. Mill and G.E. Moore.

OCU will be defended by examination of six influential objections to various of its hypotheses. In each case, thorough conceptual analysis, aided by consultation of relevant scientific facts about human nature, will reveal that the objection is seriously flawed. In the process of dispatching these negative considerations, a comprehensive positive ethical theory will emerge. Most other currently popular ethical theories are not comprehensive, but instead take no position (or several, which amounts to the same thing) on one or more of the major issues that any comprehensive ethical theory must deal with. OCU thus emerges as one of only a very few contemporary comprehensive ethical theories - and on balance the most plausible of the lot.

TABLE OF CONTENTS

Abstract	ii
Table of Contents	iii
Acknowledgement	v
Foreword	vi
 Chapter One	 Initial Considerations
1.	Introduction 1
2.	Moral entities 7
3.	Qualons and indexons 12
 Chapter Two	 The property of goodness
1.	Aristotle's dilemma 20
2.	Goodness and coloredness 25
3.	Skepticism about determinable properties 33
4.	Some puzzles resolved 42
 Chapter Three	 The concept of goodness
1.	The metaphysics of concepts of determinables 47
2.	Entertaining concepts of determinables 60
3.	The feel of goodness and the sense of right and wrong 86
4.	The open question argument 97

Chapter Four	Components of utilitarianism	
1.	Compatibilist utilitarianism	110
2.	Utilitarian rules	123
Chapter Five	Supererogation	
1.	Small beneficial sacrifices	148
2.	Valuable but not obligatory	153
3.	The sacrifice/gain ratio	180
Chapter Six	Two metaphysical background issues	
1.	Goodness and pluralism	189
2.	The Synthetic A Priori	208
3.	Conclusion	226
Bibliography		227

ACKNOWLEDGEMENT

My thanks to Randy Reiffer, John Stewart, Dale Beyerstein, Kurt Preinsperg, Leo Paquin, Paul Russell, Richard Sikora, Mark Dickson, Earl Winkler, J.L. Wisenthal, Thomas Hurka, and others for helpful discussions and useful comments on earlier drafts.

FOREWORD

The bulk of Chapter Two has been previously published; the bibliographic information is as follows.

Sleigh, Nicholas (1992), 'Objective Goodness and Aristotle's Dilemma', Journal of Value Inquiry 26, 341-51.

Chapter One

INITIAL CONSIDERATIONS

1. Introduction

In this thesis I shall defend a version of utilitarianism called 'objectivist compatibilist utilitarianism', or 'OCU'. On this version, utilitarian metaethics includes the propositions that there is an objective, intrinsic property of goodness entifying, motivating, and grounding ethics, and that act and rule utilitarianism are compatible since under plausible interpretations both true. While this metaethical theory has perhaps not been stated explicitly before, theories in this vein have been popular since the mid-19th century with Mill (1861). This branch of ethics reached the height (so far) of its influence in the early 20th century with the work of Moore (1903, 1912, 1922, 1959). I shall not defend this historical claim of antecedency, though I shall refer to some points made by these moral philosophers.

Current skepticism (among metaethicists, at least) about OCU is due largely to six objections, which this thesis will examine.

The first objection is that the proposition that an objective property of goodness exists entails a lethal dilemma:

if (a) whatever (intrinsic) goodness there is in the world exists as a property which supervenes on the natures of certain things, e.g., pleasure or honor, then the things' natures seem deprived of any direct role in explaining the goodness of things, which seems wrong. Yet if (b) whatever goodness exists does not supervene on the natures but exists wholly within them, then since some of the natures have, qua good things, nothing in common, a fortiori the natures have no property of goodness in common. If goodness is a property then either (a) or (b), hence no property of goodness exists.

Call this 'Aristotle's dilemma' after Aristotle's expression of it (1985, p. 11). Aristotle's dilemma grows from the feeling that we should be able to say more about the property of goodness than merely that it exists; we should be able to say what kind of property it is, and explain its basic features in standard metaphysical terms. Yet, it is sometimes claimed, when we try to say more about the property of goodness we run into insuperable metaphysical difficulties, as Aristotle's dilemma illustrates.

The second objection to OCU is that even if a satisfactory metaphysical account of the property of goodness can be supplied, the phenomenological difficulty remains that most people are unable to detect such a property; the concept of goodness eludes

them despite their following whatever procedures the moral objectivist recommends for the inducement of this mental state. Indeed, some have argued that no such concept is possible.

The third objection to OCU (and utilitarianism in general) is that the most plausible and popular utilitarian doctrines are forms of either act or rule utilitarianism, but that each of these doctrines suffers a repellent objection. The doctrine that an act is right if it will result in at least as much utility as any other available act is Act Utilitarianism (AU); the doctrine that an act is right if endorsed by a rule from a set whose adoption by society will result in at least as much utility as the adoption of any other set of rules is Rule Utilitarianism (RU). AU is repellent for its apparent implication that the intuitively justified rules of commonsense morality (e.g., that promises should be kept, that truth should be told, and that innocent persons should not be punished) deserve no special favor and on any given occasion should only be followed if doing so seems likely to produce more utility than not doing so. Cynicism and social disorder seem more likely than harmonious happiness to result from taking these rules so lightly. Call this the counterintuitivity objection. RU, on the other hand, is repellent for its apparent implication that rules should be followed even on occasions when doing so would not gain any utility. A moral code that makes demands unsupported by the only intrinsic value it (overtly) posits is unconvincing. This is the rule worship objection. Thus both the chief forms of

utilitarianism are seriously flawed, hence utilitarianism including OCU should be abandoned altogether.

The fourth objection to OCU is the objection from supererogation. A supererogatory act is, to define it in as much detail as space permits, an act, beyond the call of duty, that benefits someone other than the agent, and is altruistically motivated. Saints and martyrs are extreme examples of the kinds of people to whom supererogatory acts have traditionally been attributed, though at the other end of the spectrum, very small altruisms may be supererogatory, e.g., taking a stray cat to the SPCA. In the context of OCU, a supererogatory act is an altruistically motivated act, uncalled for by any rule in the short moral code of the agent's society (i.e., beyond the call of duty, hence, in a sense I shall define, not obligatory), that both increases overall utility and increases utility for other sentient beings. The agent typically sacrifices some of its own utility to make these increases, but self-sacrifice is not essential to supererogation. The fourth objection to OCU is that it has the unworkable feature of deeming some acts supererogatory. OCU's form of supererogation can be criticised in three ways, some ways applying to all forms of supererogation, other ways applying only to some forms, including OCU's.

The first criticism is that under this definition of 'supererogation', any personal sacrifice, no matter how small, for the sake of any gain for others, no matter how great, would count as supererogatory as long as it increased utility but

wasn't called for in the short moral code of the agent's society; by the intuitions of some, that doesn't seem right (R.I. Sikora, personal communication).

The second criticism of OCU's form of supererogation is that the practice of supererogation is unjustified because the concept of supererogation is paradoxical, in that a supererogatory act would make things better all things considered, but is not obligatory. This nonobligatoriness seems inconsistent with

the resilient intuition that if an action is good, this gives us a reason to do it, and if it is the best available, we have more reason to do it than to do anything else and so ought to do it. If this intuition is sound, the best available action will always be the right one, the one that one ought to do.

(Dancy 1988, p. 176)

This general criticism of any morality's form of supererogation is starkly applicable to OCU's, for according to OCU all and only utilities intrinsically ground reasons for action.

The third criticism of OCU's form of supererogation is that it is impossible that the set of rules forming OCU's short moral code could ensure that the correct sacrifice/gain ratio is on average achieved by everybody, simply because there is no such thing as a sacrifice/gain ratio that is correct for everybody.

These three criticisms jointly form the fourth objection to OCU, the objection from supererogation.

The fifth objection to OCU is that moral pluralism might be true, but that OCU is a form of moral monism, hence is implausible to the degree that moral pluralism is a live option.

The sixth objection to OCU is that OCU assumes the existence of some synthetic a priori propositions, but that no such propositions exist. OCU does make this assumption, so must defend the synthetic a priori. OCU makes this assumption by assuming that determinable properties exist; it will be seen that determinable properties, whether construed as supervening on their determinates or as existing within them as parts, stand in metaphysically necessary relations with their determinates, or with those parts of their determinates that are ontologically independent, and that the propositions describing these states of affairs are synthetic a priori. OCU also assumes the existence of analytic a priori propositions, but these are much less controversial entities, so will not be defended in this thesis.

In this thesis I shall refute these six objections to OCU, hence show that OCU is a much stronger comprehensive moral theory than is often supposed. The counterarguments I make will reveal a substantial positive moral theory. The first objection to OCU will be dealt with in ch. 2, the second objection in ch. 3, the third in ch. 4, the fourth in ch. 5, the fifth in the first subsection of ch. 6, and the sixth in the second subsection of ch. 6.

Many other objections that have at times been influential against OCUish utilitarianism have already been refuted, e.g.,

the lack of calculation time objection (refuted by Mill 1861, p. 30-32), the cold, calculating mentality objection (refuted by J.J.C. Smart 1973, p. 44-45), the objection of obsession with pleasure (refuted by Blake 1967, p. 432-33), the objection that pleasure cannot be quantified (refuted by Blake 1967, p. 434-35; see also §6.1), the objection that utilitarianism should but cannot attribute intrinsic value to the distribution of goods according to desert (refuted by Sikora, forthcoming), and the objection that utilitarianism should but cannot promote, for at least extrinsic reasons, distributions of goods according to egalitarian considerations (refuted by Brandt 1979, p. 219-20, ch. 16). Similar refutations appear elsewhere in works too numerous to mention. Some of these refutations point out proximate errors in the relevant objections, others consist in showing that the objection applies to any form of consequentialism, and appealing to the extreme implausibility of any ethical theory on which the consequences of an act are of not the slightest importance in deciding whether one should do the act. Many of these objections are susceptible to both sorts of refutation.

In the remainder of this chapter I shall explain a few details of the nature of OCU, and a metaphysical distinction relevant to my defense of OCU.

2. Moral entities

Moral language includes many kinds of terms, some referring to ethically fundamental entities, others functioning otherwise. Among the latter, and resembling the former, are terms referring to ethically derivative entities. For instance, the terms 'right', 'ought to', 'obligatory', 'permissible', and 'duty' all have ethical content but are all derived, according to OCU, from the term 'good', which in the relevant sense attributes the irreducibly ethical property of intrinsic goodness. Goodness, and its negative twin (or negative half), badness, together make possible definitions of the other ethical terms listed: e.g., in simplified form, for 'the right act', read 'the act that would do the most good'; for 'ought to do X' read 'would do the most good by doing X'; for 'is obligatory under any circumstances' read 'would uniquely do the most good'; for 'is obligatory all else equal' read 'all else equal will do the most good'; for 'is permissible' read 'would not all else equal do the most good not to'; for 'has a duty to do X' read 'all else equal would do the most good by doing X'. These definitions in some cases reflect actual usages from the natural language English, and in others are reforming definitions designed to pick out aspects of moral reality more closely than previous definitions. These sample definitions are simplified for brevity, and in full would take account of many complications, such as the right act being strictly speaking not the act that does the most good, but the one that brings about the optimal balance of good over bad, taking into account both the absolute quantities of good and bad,

and their mutual proportions (the literature on such complications is extensive; e.g., Moore 1912, ch. 2, 3). However, it is clear in what sense these full definitions would be derived from ethically fundamental entities. Since all other moral definitions are derived from goodness and badness, and since the metaphysics of badness is in every respect either the same as or the complement of the metaphysics of goodness, I shall explain and defend the entities and language of OCU almost entirely in terms of goodness. Exceptions occur where a relevant claim by another writer has been couched in terms of derivative moral entities, and it is simpler to discuss the claim in these same derivative terms than to reverse the derivations.

It might be objected to my free use of existing moral terminology, including the claim that at least some of these terms have existing senses that pick out real moral entities, that existing moral terms are hopelessly vague and ambiguous, hence that reliance on the linguistic intuitions they embody can yield no definite results in moral inquiry (e.g., Brandt 1979, p. 4-6). It must be admitted that few or no existing moral terms of ordinary language have single, clear, language-wide meanings. However, when I claim to use existing moral terminology, I mean merely that the terms in question have, among their various senses, well-established senses that pick out real entities. These well-established senses can be used to yield definite results in moral inquiry. These senses are well-established in that for many years they have been known to and used by many

people, not all of whom are extreme specialists in the relevant subject.

Terms throughout language have senses well-established in this way; for instance, a nautical term might have a sense known to almost no one but sailors, but which was a common term among this group, with a clear working sense known to almost all of such people. This is not to say that sailors could give a polished technical definition of the sense, merely that they successfully use it in their trade and easily explain it to newcomers. As Sikora remarks,

the best test I know of for whether teaching the meaning of a given sort of expression is possible is whether people can be taught by the procedure in question to use the new word with substantial agreement in regard to an open class of cases. (1981, p. 448)

The moral terms I have in mind pass this test. If words that meet this criterion, but not a stronger one, are hopelessly vague and ambiguous, making them unfit for philosophical use, then a great deal of what was hitherto taken to be excellent philosophy must be scrapped. But it is silly to force impossibly high standards upon language. When I claim that 'good' has an existing sense that I wish to use, I think this claim verified by the fact that 'good' has for centuries been used by philosophers, theologians, and many others in a way quite as practical and precise as the way in which sailors use 'marlinspike', and which

received expert clarification and endorsation almost a century ago with Moore (1903, ch. 1). Such use, clarification, and endorsation is no guarantee that the entity the sense picks out exists. I rely on the linguistic intuitions this sense embodies mostly to help suggest to the reader the sort of entity I have in mind, not to prove that such an entity exists; it is the task of this thesis to help provide evidence that the sense picks out an existing entity. The existence of the relevant sense is some evidence, but far from conclusive. But as Brandt agrees (1979, p. 7), nothing much hangs on such linguistic questions. If no suitable sense of the term 'good' had existed in English, I would have invented one.

As to my use of existing moral terms other than 'good' to help introduce topics or frame questions, even though I freely admit that such terms are in some respects vague or inaccurate, I hold that such terms are nonetheless useful in directing the reader's attention at least to the general area of inquiry I wish to pursue. Brandt is in no position to criticize this tactic, for he uses it himself: he opens his book thus:

Moral philosophy has traditionally been a systematic attempt to answer some questions of apparently universal interest about: what is worth wanting or working for; what is the best thing for an agent to do from his own point of view; what is morally right; and what is morally just. (1979, p. 1)

The tactic is endorsed by Russell:

I do not pretend to start with precise questions. I do not think you can start with anything precise. You have to achieve such precision as you can, as you go along.

(1986, p. 170)

One of the trials of conducting philosophical inquiry in a natural language is the scarcity of accurate terminology with which to begin the inquiry. However, the only alternatives are to use an artificial perfect language, which does not exist, and if it did would render much inquiry superfluous, or to conduct inquiry without using language, an activity which cannot be coherently imagined. So let's be practical and get on with the job.

3. Qualons and indexons

Elucidating certain relevant metaphysical points will require us to distinguish between two kinds of property; call them 'qualitative properties' and 'indexical properties'. Qualitative properties ontically express what a thing is, what kind of concrete stuff it is made of; indexical properties ontically express where a thing is. A qualitative property of my fridge is that it has mass; two indexical properties of my fridge are that it is to the left of the stove and that it is twice as

tall as the stove. A qualitative property of my visual sense-datum of my fridge is that it is green; two indexical properties of my visual sense-datum of my fridge are that it is to the left of my visual sense-datum of the stove and that it is twice as tall as my visual sense-datum of the stove. Thus a given qualitative property cannot be realized both mentally and physically, whereas at least some indexical properties can be realized both mentally or physically (see McDowell 1985, p. 115-16).

I call the first sort of property 'qualitative' because of the view that for things to have concrete natures is for substances to be qualified. I take 'substance' in a broadly Spinozan sense (Spinoza 1967; Bennett 1984) in which substance is the space-time continuum augmented to incorporate any other irreducible locational field. Qualitative properties are all those concrete properties possibly inhering in substance so construed. Qualitative properties may be either mental or physical (or, counterfactually, of some other sort of concrete stuff), and are necessary though not always sufficient for causal events.

I call the second sort of property 'indexical' for two etymological reasons. First, the locational fields they navigate serve as indexes to individuate whatever qualitatively indistinguishable entities they contain (hence may be termed 'indexical fields'). Two identical golden balls in an otherwise empty universe could only be distinguished by their different

spatiotemporal locations. The index reads 'ball A is at such-and-such a location, ball B is at such-and-such a location'; since these locations differ, balls A and B are different balls. Second, many such properties make essential reference to indexicals such as 'this', 'I', 'my', 'here', and 'now' (Perry 1979). For example, to have the indexical property of being to the left of the stove is to be to my left of this stove. To have the indexical property of being a meter long is to be the same length as this stick. Perhaps all indexicals reduce to some subset, perhaps not. The parts of an indexical field are identical, but can be picked out by ostension or qualitative contents. The possibility of an indexical field entails that sufficient indexical properties exist to ontically express the field, whether or not such properties are actually instantiated.

Some properties which might seem qualitative, such as the property of being circular, are indexical; in the present case, to say that a doughnut is circular is to say that to say that its doughnut parts are spatially located in a certain way relative to each other, and space is an indexical field. The notion that such properties are qualitative makes sense only if one takes the substances they inhere in to be bare particulars, which are merely conventional entities, rather than parts of a Spinozan field, which is an objective existent. To hold that a baseball is a set of properties clothing a single particular, as the bare particular theory suggests, is implausible. If I cut the baseball in half, so that there are now two objects hence two

bare particulars, at the severance of which atomic connection did the bare particular bifurcate? All the problems of identity over time in Greek ships (e.g., Chisholm 1969) attach to the bare particular theory, whereas the Spinozan substance theory is impervious to such objections because Spinozan substance is extended.

The metaphysics of qualitative and indexical properties is complex and difficult; a full defence of my interpretations of them must for brevity be omitted from this work. However, the general distinction involved is almost universally made in some guise or other. Locke, for instance, distinguishes 'primary qualities' and 'secondary qualities' in roughly such a way as to pick out, respectively, indexical and qualitative properties (1894, p. 170); Williams terms qualitative properties 'qualities' and reduces indexical properties to information "shown" in suitable statements in a logically improved language lacking predicative terms signifying the sizes, shapes, or movements of things (1969, p. 312); Husserl distinguishes pure essence, having content, from form, which has no content (1970, p. 455). The exact way in which this distinction is made is unimportant to OCU; it is important only that the distinction be made. I therefore invite those who disagree with my interpretation of the distinction to read on forbearingly, and to later mentally rearrange the relevant details to suit themselves. Indeed, I leave it open that some or all indexical properties reduce away as Williams suggests, for ontic parsimony is as important as

realism in giving an account of the phenomena underlying the distinction. Such a reduction would be consistent with realism about indexical phenomena, for indexical facts would be preserved in the states of affairs to which indexical properties were in this way reduced. Facts are truths, and by the correspondence theory of truth any truth must correspond to some real state of affairs. Not all reductions would be thus acceptable. A reduction of mental states to physical states would be in this way disfavored, for something of the conceptual flavor or character of mental states is utterly lacking in physical states, whereas the suggested reduction of indexical properties to information "shown" in suitable statements in a logically improved language lacking predicative terms signifying the sizes, shapes, or movements of things would preserve the conceptual flavor involved; the disagreement would be about a theoretical entity. The reduction of molecules to atoms is a successful example of the latter sort.

Note that the 'reduction', if the term is appropriate, of goodness to a property would not, as some people seem to think, be disfavored on the ground that the conceptual flavor or character of ethical thinking would be utterly lacking in a theoretical entity from technical metaphysics such as a property. On the contrary, to assert that a property of goodness exists would be to assert that the conceptual flavor or character involved is a real, tangible thing in the world, just as the positing of mental properties asserts the existence of the flavor

or character of the mental as a real, tangible thing in the world, not to be explained away as having its existence solely in some other real, tangible thing of a different conceptual flavor or character.

Analogously, it is sometimes objected that free will cannot have a mechanistic explanation, since the mechanistic displaces the purposive, yet free will involves purpose (see Dennett 1982). If events of free choice are reduced to mechanistic sequences, says the objection, then the responsible character of freedom will be lost. The vague intuition underlying the objection seems to be that if freedom is precipitated out as concrete objects linked causally, then it won't be ethereal or fluid enough to match our feelings when we do things freely. To this objection the modern response is that freedom and causation are compatible, perhaps even necessarily linked (see Hobart 1934). Ethereal, fluid phenomena are nonetheless concrete objects capable of standing in causal relations with other objects. As with freedom, so with goodness; ethically ethereal goodness can be identified as a qualitative hence concrete property without losing anything of its character.

Define an object composed wholly of qualitative properties as a 'qualon'; the color set green and white is a qualon. Define an object composed wholly of indexical properties as an 'indexon'; a rectangle is an indexon. 'Quale' and 'qualia' would be just as good as 'qualon' and 'qualons' for this purpose, but the former terms already have established usages according to

which the subjective, conscious, experiential parts of thinking brains are termed 'qualia' (e.g., Shoemaker 1990, p. 110; Angel 1989, p. 62). If one used the term 'qualia' in this sense to pick out the items I call 'qualons', then one would reject the possibility of being conscious of indexical properties; however, we are conscious of such properties. Also, some non-indexical properties are physical, which 'qualon' allows for, whereas qualia are always mental. Rather than confuse matters by coining a new sense for 'qualia' taking care of these objections, it seems simpler to coin 'qualon'.

Part of the aim of distinguishing qualons and indexons is to help dispel the widespread confusion between scientific inquiry and inquiry about the physical nature of reality. Scientific inquiry is empirical inquiry about instantiation patterns among properties, chiefly about causal property successions. The comparative ease of scientific inquiry about purely physical property instantiation patterns, and difficulty of scientific inquiry about sets of properties some of whose members are mental, has caused many to believe that science is concerned only with physical properties (e.g., R.N. Smart 1964, p. 106-07), or worse, that mental properties are somehow less than fully real, and are only intellectually respectable when reduced to physical properties. Such confusion has sometimes led to unjustified moral skepticism (see Taylor 1989, p. 235-36).

It is unclear whether my interpretation of mental and physical entities is a form of monism or dualism. As will be

discussed in §3.1, the monism/dualism distinction is difficult to construe in the traditional way, i.e., exclusively. My interpretation of mental and physical entities holds that mental and physical concrete entities each constitute quite real, tangible kinds of thing in the world, having distinct conceptual flavors or characters, but that both mental and physical concrete entities are qualons. It is unclear why one should pay more attention to their difference than their similarity, or vice versa, and that is what one needs to do in order to opt for monism or dualism construed exclusively.

Chapter Two

THE PROPERTY OF GOODNESS

1. Aristotle's dilemma

The first objection to OCU is that one of its premises, the proposition that a property of goodness exists, in a robust, objective, intrinsic sense in the spirit of Moore (1903, ch. 1; 1922, ch. 8; 1959), is subject to a lethal dilemma:

if (a) whatever (intrinsic) goodness there is in the world exists as a property which supervenes on the natures of certain things, e.g., pleasure or honor, then the things' natures seem deprived of any direct role in explaining the goodness of things, which seems wrong. Yet if (b) whatever goodness exists does not supervene on the natures but exists wholly within them, then since some of the natures have, qua good things, nothing in common, a fortiori the natures have no property of goodness in common. If goodness is a property then either (a) or (b), hence no property of goodness exists.

Call this 'Aristotle's dilemma' after Aristotle's expression of it (1985, p. 11; see also Strawson 1949, p. 26-27). Aristotle's dilemma grows from the feeling that we should be able to say more about the property of goodness than merely that it exists; we should be able to say what kind of property it is, and explain its basic features in standard metaphysical terms. Yet, it is sometimes claimed, when we try to say more about the property of goodness we run into insuperable metaphysical difficulties, as Aristotle's dilemma illustrates. Some writers are pessimistic about the prospects of answering such difficulties; for instance, J.L. Mackie says that

if there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe. . . . What is the connection between the natural fact that an action is a piece of deliberate cruelty - say, causing pain just for fun - and the moral fact that it is wrong? It cannot be an entailment, a logical or semantic necessity. Yet it is not merely that the two features occur together.

(1977, p. 38)

A Moorean moral objectivist (as, for brevity, I call the proponent of a Moorean property of intrinsic goodness, though as Moore notes, there is also a sense in which one can be a moral

objectivist without positing intrinsic goodness or any other intrinsically valuable ethical entity - 1922, p. 255; various versions of moral objectivism will be formally distinguished in §3.2) must answer such questions, both by giving a positive general account of moral entities and by showing how this account resolves standing particular objections. I think that Moorean moral objectivism can supply such answers, and as evidence show how it, or at least the version of it developed in this thesis, can deal with one of the most influential metaphysical objections, Aristotle's dilemma.

I argue that the Moorean moral objectivist can escape Aristotle's dilemma through either fork (a) or fork (b), depending on how a metaphysical issue is resolved. The solution is that the relevant relation between the property of goodness and the properties comprising the natures of good things is a species of entailment. To say that property A entails property B is short for saying that the proposition that x has A entails the proposition that x has B. This conditional is itself true in virtue of the natures of A and B in a way I shall discuss.

I interpret supervenience to obtain only between ontologically independent entities, so that a complex property is not supervened on by its parts. I use 'entailment' not in the narrow sense of strict logical implication, but in the broader sense of conceptual necessitation. While entailment in this broader sense is a relation whose instances are knowable a priori, propositions asserting these instances may be either

analytic or synthetic. I take a proposition to be analytic if its logical form, as revealed by the meanings or meaningful uses of its terms, reveals its truth-value, and synthetic otherwise; I take a priori knowledge of a proposition's truth-value to rest essentially on a correct understanding of the meanings or meaningful uses of the proposition's terms.

Thus there may be different kinds of entailment. To escape Aristotle's dilemma, the Moorean moral objectivist specifies that moral entailment is of the same general kind as the entailment by which necessarily a red thing is colored. Perhaps the resemblance between color entailment and moral entailment (as I call these two kinds of entailment) is not perfect, but the Moorean moral objectivist need only hold that the two kinds of entailment are alike in whatever common aspect of their structures allows them to escape dilemmas of the form of Aristotle's. Such an aspect exists (at least in color entailment) because, given that coloredness exists, Aristotle's dilemma cannot be used to prove the nonexistence of the property of coloredness.

(Using the term 'common aspect' is to some extent question-begging, but like Neurath's boat-carpenter, I can do little about it beyond noting that all rational inquiries suffer some form of this defect, yet many are successful - see Neurath 1937, p. 276. We must frame our questions from the presupposition-laden linguistic material we find at hand, hoping that ensuing inquiry will, as often happens, reveal or clarify

the presuppositions and shed light on their truth-values. These presuppositions must eventually be analysed and perhaps discarded; their preliminary use for pragmatic reasons is not an excuse for intellectual mush, such as Rorty 1989).

Thus I follow the time-honored strategy of seeking out respectable siblings for a shady proposed entity. Goodness resembles other entities in various ways. Goodness and its negative twin (or negative half), badness, together make up the whole of an intensity continuum, just as do the properties of heat and coldness. Goodness and humans, like pleasure and humans, are such that necessarily humans who grasp the concept of goodness (or pleasure) have some positive disposition toward goodness (or pleasure), though not necessarily to the exclusion of a simultaneous negative disposition toward goodness (or pleasure). And goodness and coloredness are alike in that they are determinable properties, by which I mean (at least) that they cannot be instantiated in the absence of the instantiation of one of a range of other, determinate properties, such as pleasure and redness, in virtue of whose instantiation the determinable property in question is also instantiated.

My defense of this solution to Aristotle's dilemma has two parts: first, an explanation of how I interpret coloredness, and how the property thus interpreted, hence also goodness, escapes Aristotle's dilemma; and second, a discussion of some metaphysical objections to the existence of properties such as goodness and coloredness which ontically express the unity of

ranges of less controversial properties.

2. Goodness and coloredness

I interpret coloredness in such a way as to avoid some relevant disanalogies between goodness and coloredness as the latter is sometimes interpreted. The first disanalogy stems from the view that determinate colors are dyadic properties, so that what is, say, red to me may be green or transparent to you.

By a dyadic property I mean a two-place relation. This interpretation views relations and properties as not fundamentally different, but merely as universals obtaining of different numbers of particulars. Armstrong attributes this view to Plato and possibly Aristotle (1989, p. 15-17), and Russell (1912, p. 94-96) and Oddie (1991, p. 21) subscribe to this view. To illustrate, the dyadic property of heaviness obtains of the ordered set {a 5 year old child, a sack of potatoes}, and the dyadic property of lightness obtains of the ordered set {the world's most muscular human, a sack of potatoes}, for the sack of potatoes is light to the one human trying to lift it, and heavy to the other. The sack of potatoes thus participates in both lightness and heaviness, objectively but relatively. A monadic property is a 'one-place relation', i.e., a property in the non-Russellian sense in which the word 'property' is often used. Colors such as redness and greenness are sometimes interpreted as

objective but dyadic properties, e.g., by Dolby 1973, so that the same physical object might be red to me but green to you. On this interpretation an object might to you be green, hence colored, and to me be transparent, hence not colored.

Plainly, the Moorean moral objectivist cannot allow the analogue for goods and goodness that one person might correctly judge a state of affairs to be pleasant, hence morally good, and another person correctly judge the identical state of affairs not to be pleasant, hence not morally good. Such relativism would defeat the whole spirit of Moorean metaethics (also the letter; e.g., Moore 1959, p. 96-97).

Thus life would be easier for the Moorean moral objectivist if colors were non-relative, monadic properties. If there were no disanalogy between goods and colors, hence also, *mutatis mutandis*, between goodness and coloredness, in this respect, a *fortiori* no disanalogy would be relevant to Aristotle's dilemma. As it happens, color terms such as 'red' and 'green' are used in several senses in ordinary English, and in one sense they are monadic. I intend this latter sense by 'red' and 'green', and so *mutatis mutandis* for other color terms such as 'coloredness'. Clearly colors in this sense exist.

Consider a person who has normal human vision and sees a ripe tomato. The existence of the vivid phenomenal quality redness such a one experiences does not entail the existence of any ripe tomato or light-beam. The entity in which the vivid phenomenal quality is instantiated, namely the sense-datum, or

seeing, or perception, or piece of visual field, or appearance, or whatever you want to call it, could exist non-veridically in the absence of any ripe tomato or light-beam, as is shown by the illusions we all experience from time to time, and by the color patches we sometimes experience which aren't of any represented thing outside our minds. If two qualitative properties are ever ontologically independent then they are always ontologically (if not causally) independent. Thus, if a tomato and an instance of my having a sense-datum (or whatever) depicting such a tomato happen to coincide, it is an accident of history. The vivid phenomenal quality redness is ontologically independent of the non-mental objects the color relativist must claim it ontologically depends on. Perhaps relative, dyadic colors exist also; if so they seem dispositional and reducible at least in part to non-mental properties of physical objects.

A second disanalogy between coloredness and goodness is that while a clear distinction exists between intrinsic and extrinsic goodness, to speak of something's being intrinsically rather than extrinsically colored is senseless. However, construed as an objection this claim trades upon the ambiguity of the term 'senseless'. Certainly to speak of something's being extrinsically rather than intrinsically colored is senseless if this is taken to express the view that the term 'extrinsically colored' is assigned no sense in the natural language English. The term has no meaning in ordinary speech. However, the claim is implausible if taken to express the view that to apply the

intrinsic-extrinsic distinction to coloredness would be futile because the result would be nonsense. On the contrary, we could easily coin a sense for 'extrinsic coloredness' analogous to the sense of 'extrinsic goodness'. We could hold that just as an extrinsically good thing, such as a cold beer on a hot day, is by definition an important cause of an intrinsically good thing, such as a pleasant, exhilarating taste sensation, an extrinsically colored thing, such as the firing of an optic nerve, is by definition an important cause of an intrinsically colored thing, such as a visual sense-datum. Few would find it useful or interesting to talk about such extrinsically colored things, which explains the absence of 'extrinsically colored' from ordinary English, but such talk would not be incoherent. Thus the intrinsic-extrinsic distinction grounds no relevant disanalogy between coloredness and goodness.

Coloredness thus interpreted escapes Aristotle's dilemma through either fork (a) or fork (b). Consider these forks in relation to coloredness:

if (a) whatever (intrinsic) coloredness there is in the world exists as a property which supervenes on the natures of certain things, e.g., red things or yellow things, then the things' natures seem deprived of any direct role in explaining the coloredness of things, which seems wrong. Yet if (b) whatever coloredness exists

does not supervene on the natures but exists wholly within them, then since some of the natures have, qua colored things, nothing in common, a fortiori the natures have no property of coloredness in common. If coloredness is a property then either (a) or (b), hence no property of coloredness exists.

Which fork coloredness escapes through depends on whether color entailment statements are analytic or synthetic. If color (and by hypothesis hence moral) entailment statements are analytic, then fork (b) is indicated; if synthetic, fork (a). Let us see how this works.

If color entailment propositions are analytic, then the proposition 'All red things are colored' is of the form $x((Fx.Cx) \rightarrow Cx)$ where 'C' attributes coloredness and 'F' attributes some property F which combined with coloredness makes redness. One might object that a property such as F is inconceivable. A Moorean moral objectivist might attribute this inconceivability to human cognitive deficiency, or, alternatively, admit the epistemic possibility that necessarily F cannot be conceived, but deny that the truth of this possibility entails that no such property exists. Instead, the Moorean moral objectivist might appeal to Bennett's notion of unabtractable differentiae:

There are in fact unabtractable
differentiae, or so it seems. If we ask,

'What is the value of F such that redness is colouredness-and-F?' there seems to be no answer (except the trivializing one 'It is the property of being either not coloured or else red'). . . . we must drop the idea that there is a concept corresponding to every property. (1984, p. 144-45)

Husserl seems to agree (1970, p. 454). If a range of properties such as F exists, then color entailment statements are analytic, and coloredness is a part of redness, yellowness, etc., thus satisfying the requirement of fork (b) of Aristotle's dilemma that coloredness not supervene on the natures of red things, yellow things, etc., but exist wholly within them.

Simultaneously falsified is the dilemma's assumption that some of the natures of red things, yellow things, etc. have, qua colored things, nothing in common. These things have in their natures coloredness in common. Thus fork (b) is satisfied yet its troublesome alleged consequence is avoided.

Likewise, the Moorean moral objectivist might hold that a range of unabstractable differentiae exists so that each member of the range combined with goodness forms a good, such as pleasure. Thus the impossibility of factoring out the concept of a good such as pleasure into two concepts, one being the concept of goodness, is accounted for without giving up the view that the property of pleasure (or whatever) factors into two properties, one being goodness. Fork (b) is satisfied, since goodness exists

wholly within the natures of good things, and the assumption is avoided that these things have, qua good things, nothing in common. They have goodness itself in common.

One might object that if no concepts correspond to unabstractable differentiae such as F , then a statement of the form $x((Fx.Cx) \rightarrow Cx)$ cannot express an entailment, since entailment is conceptual necessitation. However, while F cannot be conceived, the form $x((Fx.Cx) \rightarrow Cx)$ can be conceived, and this form renders the statements in question conceptually necessary, hence entailments.

Nor does it help to object that coloredness might turn out to be not a new sui generis property, but a familiar property, such as being a light-emission of some minimal wavelength. It is not crucial to Aristotle's dilemma that the determinable bear no internal relation to anything save (each of) its determinates. What is crucial is that the determinable bear some suitable internal relation to its determinates. The fact that the property of being a light-emission of some minimal wave-length apparently lacks a suitable internal relation to redness and yellowness suggests that it is not the common part sought. Likewise, Aristotle's dilemma would still be satisfied if objective goodness turned out to be cheddar cheese, so long as the relevant internal relations obtained (though since they don't, it isn't).

If, on the other hand, color entailment statements are synthetic, then the statement 'All red things are colored' is of

the form $x(Rx \rightarrow Cx)$ where 'R' attributes redness, 'C' attributes coloredness, and 'Rx' is not further analyzable as 'Bx.Cx'. This satisfies the requirement of fork (a) that coloredness exist not within the natures of red things, yellow things, etc., but separately supervening on these natures. Since the truth of 'All red things are colored' is conceptually necessary, a thing's redness does play a direct role in explaining that thing's coloredness. If the thing were not red but otherwise exactly as it is, we should have no explanation for its being colored. A significantly useful explanation must be of one thing as due to another, and conceptual necessitation is an extremely strong, ontologically rather than linguistically based relation; hence, the redness of a thing's explaining its coloredness is as about as strong and direct as an explanation can get. Thus the unwanted alleged consequence of the requirement of fork (a) is avoided.

Similarly, the Moorean moral objectivist might view goodness and any particular good as separate entities, thus satisfying the requirement of fork (a) that goodness exist not within the natures of particular goods but as a property supervening upon them. This view is consistent with the Moorean moral objectivist's further possible view that moral entailment statements such as 'All pleasures are good' are, as synthetic a priori truths, conceptually necessary, so that the pleasantness of a thing explains its goodness about as strongly and directly as anything can be explained. Thus the unwanted alleged

consequence of the requirement of fork (a) is avoided.

The Moorean moral objectivist can escape Aristotle's dilemma without being committed to either fork (a) or fork (b) alone. Moorean moral objectivism is secure so long as goods and goodness are regarded as metaphysically analogous to colors and coloredness, leaving for another day the question whether color entailment statements and moral statements are analytic or synthetic.

3. Skepticism about determinable properties

A critic of this solution to Aristotle's dilemma for objective goodness might allow that goodness and coloredness are metaphysically analogous in the required way but hold that I have removed little of the mystery they are cloaked in and have merely provided a reason for doubting the existence of a property of coloredness as well as a property of goodness. I reply that while determinable properties are somewhat mysterious, their absence would be more mysterious. Determinables account for resemblances between determinates. If determinables don't exist, then either determinates don't resemble each other, or their resemblances exist but have other accounts.

The first disjunct is false. Determinate properties do resemble each other in various ways. In particular, every color property resembles every other color property in being a color.

Armstrong is right that "these recalcitrant resemblances seem objective phenomena, demanding an ontological analysis" (1978, vol. 2, p. 116; on some other points regarding resemblance I paraphrase Armstrong).

The second disjunct, that these resemblances exist but have other accounts, is dubious because both the chief rival accounts are dubious. To show this I review the five chief ontological accounts of resemblance. Three turn out to posit determinables (including the two fairly plausible accounts already examined) - hence they are benign competitors for purposes of the moral objectivist - and the other two are flawed.

We have examined the first account of resemblance, that resembling determinates do so in virtue of having common parts - call this the 'common part theory'. We may not be directly aware of the complexity and partial identity of, say, the colors, but we can hold that they are ontologically complex though epistemologically simple. Given that the colors contain a common part ontically expressing their color resemblance, clearly the common part is itself a property (a definite way of being, a one-place universal), and 'coloredness' a useful name for it. If the colors are epistemologically simple, then their unshared parts are unabstractable differentiae, perhaps complex. Coloredness thus construed is a determinable property, escaping Aristotle's dilemma via fork (b); likewise, says the Moorean moral objectivist, with goodness.

We have also examined the second account of resemblance,

that two objects resembling each other instantiate the same determinable, each in virtue of its also instantiating a member of a range of ontologically independent determinates. The determinable is not part of but supervenes on the determinates, and statements asserting this supervenience are synthetic a priori. This account allows determinables to escape Aristotle's dilemma via fork (a), and benefits from whatever plausibility attaches to the notion of synthetic a priori propositions - which is considerable (see §6.2, and, e.g., Sikora 1981). Thus while this account of resemblance competes with the common part theory, it is a benign competitor for purposes of the Moorean moral objectivist.

The third account of resemblance is that determinables underwrite resemblance but are second-order properties: common properties of the universals concerned. Bertrand Russell favors this view: "I should regard 'red is a color' as a genuine subject-predicate proposition, assigning to the substance red the quality color" (1959, p. 171). We need not debate the merit of this account, since if false it poses no threat to the previous two benign and plausible accounts of resemblance, and if true it lets the second-order property of coloredness escape Aristotle's dilemma via fork (a). Let's see how this latter claim works.

The two chief conditions for fork (a) are that: first, redness and coloredness (for example) be ontologically independent entities, so that the latter can supervene on the former, and; second, redness and coloredness (for example) be

such that the nature of the former explains the former's instantiating the latter.

The first condition is satisfied by the third account of resemblance, since the main point of positing substances is to give properties something ontologically distinct from themselves in which to inhere. If determinate properties are taken as substances, then their determinables must be ontologically distinct. This line requires that if supervenience is defined with respect to a set of objects which if they share all properties in the supervenience base will also share all properties in the supervening set (e.g., Kim 1978, p. 152), then an object must be deemed to possess albeit indirectly the properties of its properties.

The second condition is that redness and coloredness (for example) must be such that the nature of the former explains the former's instantiating the latter. If, as it must, this third account of resemblance allows that the relevant propositions, such as that the substance red has coloredness, are knowable a priori, and since the ontological distinctness of, say, the substance red and the property coloredness makes the propositions synthetic, it follows that, say, the nature of redness does explain red's coloredness. Propositions knowable a priori may be known to be true simply from the meanings or meaningful uses of the terms involved, and since the relevant propositions cannot be known to be true from their logical forms as revealed by the meanings or meaningful uses of their terms, nothing is left to

explain the a prioricity of knowledge of their truth but the natures of the referents of their property terms - i.e., qualons rather than indexons - as revealed by the meanings or meaningful uses of these terms. These natures must therefore stand in an appropriate ontologically necessary relation, this relation grounding the explanation that, for instance, the nature of redness yields of red's coloredness. Thus the third account of resemblance posits determinable properties which satisfy both chief conditions of fork (a) and hence is benign for purposes of the Moorean moral objectivist.

The fourth account of resemblance, discussed by Johnson (1921, ch. 11), is that the unity of a class of determinates such as the shapes or colors is constituted and exhausted by the fact that co-determinates cannot simultaneously qualify the same particular. These classes are simply class-of-incompatibles.

This account is inconsistent with Moorean moral objectivism, which intends much more by 'Pleasure is good' than that pleasure is incompatible with each of a set of other things, all mutually incompatible, and perhaps does not even intend these incompatibilities. However, this account is gravely flawed.

First, it is unclear that all classes of determinates which show by resemblance an intrinsic unity form a class-of-incompatibles. Perhaps a thing can be both sweet and sour. Second, this account leaves unanswered the question what about the colors, and so on, makes them incompatible. Third, the view that co-determinates have nothing in common except their

incompatibility is phenomenologically implausible. The class of shapes or the class of colors appear to have much more in common than that. They appear to display fully and equally a common, definite way of being. Fourth, determinates that don't resemble each other at all might end up under the same determinable simply because they could not inhere in the same particular. Thus the fourth account of resemblance is so flawed that it would be very mysterious if it were true and there were no determinables.

The fourth account of resemblance might be improved, respecting the fourth objection to it, by specifying that determinates could not fall under the same determinable unless they could conceivably individually inhere in the same particular. The plausibility of this amendment depends on what counts as a particular. If experiences count as particulars and a painful experience can't conceivably be square, then painfulness and squareness would on the unamended fourth (Johnsonian) account resemble one another. However, it might be held that an experience can't conceivably be square whether it is painful or not, hence that on the amended account painfulness and squareness would not resemble each other - phenomenologically a more plausible result.

However, this account of particulars was denied in §1.3. On the Spinozan account of substance therein endorsed, the full, underived ontic account of particularity in an experience (or whatever) would be the extended region of time, and whatever other indexical fields mental events are located in, that the

experience's qualitative properties were instantiated in.

The amendment suggested above for repairing the fourth account of determinables, respecting the fourth objection to it, is on this Spinozan account of particularity dubious because it prevents determinables from forming hierarchical structures. If, e.g., the determinate reds fell under the determinable redness because they could individually but not jointly inhere in the relevant indexical locations, and other determinates such as the yellows were excluded from redness because they could not possibly inhere in these indexical locations, then the reds and yellows would be barred from falling under a common higher-level determinable such as coloredness, for they could not individually inhere in any common indexical locations. Since a hierarchy of determinables seems plausible, the amendment fails.

The fifth account of resemblance is in terms of second-order relations between the determinates concerned. This account is unhelpful to Moorean moral objectivism since pointing at the resemblance between good things is only a way of demonstrating the existence of a feature that each good would have had even had it been the only good, hence lacking relations with any other good. However, this account of resemblance suffers two objections.

First, every class of resembling determinates is distinguished not only by the resemblances between its members but also by the lack of this sort of resemblance between each member and each determinate not a member of that class.

Analogously to the proposed Moorean moral objectivist, the coloredness objectivist holds that each color would have been a color even had it been the only color. The single color would still be distinguished by a respect in which it appeared parallel to a sound or smell, but not one. Positing the determinable coloredness would account for this appearance, for it would be instantiated, whereas positing a coloredness relation would not, for it would be uninstantiated. The coloredness relationist might try to escape this objection by holding that a single color would instantiate the coloredness relation reflexively, but if that secured the single color as a color, then surely it would also do so for each of the multiplicity of actual colors. It would then be redundant to posit non-reflexive instantiations of the coloredness relation, hence there would be no reason to view that universal as dyadic rather than monadic. Monadic universals are simpler than dyadic, and parsimony favors the simpler of two entities competing for and able to do the same explanatory job.

Second, parsimony also counts against a coloredness relation instantiated non-reflexively among many determinates. Suppose that n colors form a resemblance class. Each pair of colors instantiates the coloredness relation, for each member of each pair resembles the other in the relevant way. Thus a dyadic universal is instantiated $n(n-1)/2$ times. Positing the determinable property of coloredness, however, yields a monadic universal instantiated n times. If $n > 3$, then $n(n-1)/2 > n$, hence the determinable property of coloredness is instantiated

fewer times than the coloredness relation, hence is favored by parsimony. If $n = 3$, then $n(n-1)/2 = n$, leaving parsimony to favor a monadic over a dyadic universal for the class. If $n = 2$, then we must compare a dyadic universal instantiated once with a monadic universal instantiated twice, and parsimony favors type economy over token economy, all else being equal. Since each universal does the same explanatory work, parsimony tells us in each possible case of a multiplicity of determinates to reject a coloredness relation in favor of a coloredness determinable.

A coloredness relation does not avoid the objection to a coloredness determinable that it is phenomenologically unclear that anything about redness is identical to anything about blueness; it is phenomenologically no clearer that anything about the red-green resemblance is identical to anything about the blue-orange resemblance. If coloredness relations are not taken as identical, then the unity of a class of first-order universals is being analyzed in terms of the unity of a class of second-order universals; a vicious regress follows.

Thus of the five chief accounts of determinate resemblance, three posit determinables and the remaining two are seriously flawed. Therefore, when the determinable property skeptic asserts that all determinable properties are mysterious, the Moorean moral objectivist can admit that this is to some extent so but maintain that since determinate class resemblances are objective phenomena, demanding an ontological analysis, yet all proposed accounts are partially mysterious, mystery while

unwelcome is not fatal to these accounts. All three determinable property accounts escape Aristotle's dilemma, including the least mysterious account, the common part theory. While Moorean moral objectivism in the long run owes philosophy an explanation of these remaining mysterious aspects, it is under no obligation to produce this explanation at once; no more than set theory is obliged to produce at once a transparently clear answer to Russell's paradox. Most substantive philosophical theories involve metaphysical presuppositions of lingering mystery.

4. Some puzzles resolved

The five rival accounts discussed are not epistemically mutually exclusive competitors in the competition to ontologically analyze color resemblance (hence also, if my analogy works, moral resemblance) and the other determinate class resemblances. While Occam's Razor urges us to accept only one account, it is conceivable that investigation of various determinate class resemblances might force us to accept more than one, though no more than one for any given determinate class resemblance.

A final possible disanalogy between goodness and coloredness: one might object that while colored things manifestly do lie on a continuum, good things could not lie on a continuum as is suggested (though not entailed) by the analogy

between goodness and determinables such as coloredness, because good things such as pleasure, honor, and truth are so disparate. The Moorean moral objectivist can meet this objection by accepting the constraint on moral theories that their sets of intrinsically good things not have such disparate memberships. This constraint appears to damage Moorean moral objectivism, since things as disparate as pleasure, honor, and truth are morally good yet are too disparate to lie on a continuum. However, some versions of Moorean moral objectivism accommodate both suppositions with minimal strain on our intuitions. A hedonist Moorean moral objectivist might hold that intrinsically good things are all and only pleasures, hence not too disparate, and that things such as pleasure, honor, and truth while disparate are all morally good, though some merely extrinsically morally good.

Even if intrinsically good things do not form an unbroken continuum, this disparity would not entail that they are too disparate to form a resemblance class of the required sort. They might still be fully and equally goods, just as red, yellow, and blue would be fully and equally colors even if they were the only colors. Some forms of moral pluralism are consistent with moral objectivism (see §6.1).

When Mackie asks what sort of property goodness is, and how it is connected with good-making properties, the moral objectivist has a straightforward answer: goodness is a determinable property, connected with good-making properties by

entailment.

This interpretation of goodness perhaps explains something that puzzled Moore, who says:

many . . . predicates of value . . . are
intrinsic kinds of value, . . . yet none of
them are intrinsic properties, in the sense
in which such properties as "yellow" or the
property of "being a state of pleasure" are
intrinsic properties. (1922, p. 272)

Nonetheless, such values are intrinsic in that they "share with intrinsic properties this characteristic of depending only on the intrinsic nature of what possesses them" (1922, p. 273). The difference between intrinsic predicates of value and of such properties seems to be "that intrinsic properties seem to describe the intrinsic nature of what possesses them in a sense in which predicates of value never do" (1922, p. 274). This difference is identified by the Moorean moral objectivist of OCU as the difference between determinate and determinable properties. An object may be completely specified by its determinate properties alone, in the sense that these properties plus any other properties entailed by these properties are all the properties the object possesses; yet the object's determinable properties are nonetheless intrinsic to it in that they depend on nothing else but their bearer object's intrinsic nature. Determinable properties are reactive in that they are a function of their bearer object's determinate properties;

intrinsic values are also reactive in this way, hence it is in this respect metaphysically convenient to interpret intrinsic values as determinables. Intrinsic values are thus distinguished from determinate properties, which are not reactive. This characteristic of non-reactivity would be what Moore sought when he said:

it seems to me quite obvious that . . . there must be some characteristic belonging to intrinsic properties which predicates of value never have . . . only I can't see what it is. (1922, p. 274)

Moore would not necessarily agree with this interpretation of his thought. Also, if determinates and determinables form hierarchies, so that a set of same-level determinables can serve as a range of determinates for a higher-level determinable, then this account of non-reactivity must be correspondingly amended; e.g., by indexing reactivity and non-reactivity to appropriate levels.

Interpreting intrinsic goodness as a determinable also allows the Moorean moral objectivist to reply to the objection, mentioned by Moore, that "the goodness of God . . . cannot be a quality, which he might get or lose" (1959, p. 97). Taking 'quality' to mean a monadic property, as opposed to a relation, the objection's presupposition that each of an object's properties may be freely detached is false; a determinable property can't be detached without pulling away its determinates

with it. If goodness is a determinable, then it couldn't be detached from God without pulling away the good things in God, which would be impossible given standard intellectual accounts of God's nature; e.g., Brentano: "do we not think of God as though he were the epitome of everything that is good, but raised to an infinite degree?" (1969, p. 41). To remove such a god's good characteristics would be like removing the wool from a piece of knitting.

To conclude, the Moorean moral objectivist can escape Aristotle's dilemma for a property of intrinsic goodness by comparing goodness with coloredness and interpreting the relation between goodness and good things as a species of entailment, ontologically founded on a species of determinable-determinate property relation. A tougher job for the Moorean moral objectivist is to explain why moral entailments are less easily grasped than color entailments or simple mathematical entailments (this job can now be done - see Sleigh 1993, ch. 10). The mere availability of an adequate metaphysical structure for a property does not guarantee that the property exists; otherwise a theist might assert the existence of a property of intrinsic holiness, to the utter frustration of the atheist.

The first objection to OCU is thus refuted.

Chapter Three

THE CONCEPT OF GOODNESS

1. The metaphysics of concepts of determinables

The second objection to OCU is that even if a satisfactory positive metaphysical account of the property of goodness can be supplied, as I claim to have done in ch. 2, the phenomenological difficulty remains that most people are unable to directly detect such a property; the concept of goodness eludes them despite their following whatever procedures the Moorean moral objectivist recommends for the inducement of this mental state. Indeed, some have argued that no such concept is possible. In this third chapter a positive account of the concept of goodness will be given, for this entity too is important to Moorean moral objectivism.

It is epistemically possible for a property to exist without a second entity, a concept of the property, also existing; Bennett (1984, p. 144-45) has suggested that some properties, unabstractable differentiae, may in fact have no corresponding concepts. Therefore the nonexistence of a concept of goodness would not entail the nonexistence of a property of goodness. However, the absence of such a concept would be evidence against

the existence of the posited unconceived and inconceivable object, just as the unconceivedness and inconceivability of an integer between 2 and 3 are part of our evidence against the existence of such an integer. A sufficient accumulation of negative evidence by ontic parsimony refutes the posited existence of an object. Skeptics of Moorean moral objectivism are thus warranted in probing the possible nonexistence of the concept of goodness.

Brandt, for instance, is a skeptic about concepts of objective moral properties:

although admitting that there are ethical
words, we ask the nonnaturalist: In what
sense are there concepts of . . .
unobservable, simple ethical properties . . .
corresponding to these words? . . . We do not
have concepts of these properties . . . in
the sense of ability to conjure up an image
of them, as we can in the case of colors or
shapes. (1959, p. 189-90)

In reply, the Moorean moral objectivist might allow that no concepts of moral properties exist as images, but hold that moral concepts are other, non-imaging sorts of concept, for not all concepts are images. This line must either identify the mode of moral concepts as some familiar non-imaging sort, or posit a sui generis non-imaging conceptual mode unique to ethics. The latter variant is assumed by 'ethical intuitionism', as the view has

been called that awareness of ethical facts requires
a different kind of awareness . . . neither
sensory nor introspectual . . . , which the
specialists call "intuition". . . . The
fundamental cognitive situation in morals is
that in which we intuit the rightness of a
particular action or the goodness of a
particular state of affairs. (Strawson 1949,
p. 24)

As with the property of goodness itself, the Moorean moral
objectivist must give a positive account of the disputed entity,
for any plausible metaethic must supply fundamental entities
cohering satisfactorily with our human moral phenomenology.
Though humans have much to learn in ethics, it is surely false
that after their thousands of years of intellectual effort even
the most elementary moral considerations still elude them.

My account of moral concepts is derived from my account of
moral properties. Goodness is a determinable property, so one
would expect the concept of goodness to in some ways resemble the
concepts of other determinable properties. So when Brandt asks
how we can have concepts of moral properties, when we lack the
ability to conjure up images of them, as we can in the case of
colors or shapes, the Moorean moral objectivist can reply that we
conceive moral properties as we conceive many familiar
determinable properties, abstractly rather than by conjuring up
images. The properties of coloredness and triangularity yield

concepts graspable though abstracted from any images, e.g., of particular colors and triangles; likewise with goodness and goods. This solution identifies the mode of moral concepts as a familiar non-imaging sort, the abstract mode, rather than as a sui generis conceptual mode unique to ethics (i.e., a mode unique to 'non-natural' qualities - as ethical properties have been termed by Moore 1903, p. 13-14, to little advantage, I think, for as he suggests, the important questions for Moorean moral objectivism are unaffected by the natural/non-natural distinction).

The distinctness of such concepts of determinables from images, e.g. of determinates, does not entail that the former can be entertained without the latter. Perhaps just as one cannot conceive of a colored object without a shape, even though the concepts of the color and shape are distinct, one also cannot conceive of a determinable without a determinate, hence cannot conceive of goodness without a good, such as pleasure. However, Moorean moral objectivism is not committed to this possibility.

The concept of goodness is distinct from any positive emotional reaction to the concept; as Brentano puts it,

the appearance of virtue is more agreeable than that of moral perversity. But the essential superiority of what is moral does not consist in this fact . . . ; it is, rather, a certain intrinsic correctness which, however, may also have a certain

superiority in appearance. (1969, p. 10-11)

One might hold that necessarily humans who grasp the concept of goodness have some positive disposition toward goodness, but such a one would be wise to allow also that such humans might have a simultaneous negative disposition toward goodness. The exact way in which goodness generates this positive disposition, which under some circumstances is sufficient to motivate action, is too complex to discuss here, but agrees almost completely with Nagel's plausible account of necessarily motivating objective moral truths (1970). Suffice it to say that dispositions, as reactions to items, are ontologically independent from them, though some dispositions also seem metaphysically necessarily connected to the items they are reactions to, analogously to effects being metaphysically necessarily connected to their causes (see also Goldstein 1989, p. 260). Besides this direct moral motivation, an agent may have other, indirect motivations for doing right acts, e.g., a desire to be thought helpful.

As with the property of goodness, a comparatively clear positive account of the concept of goodness suggests solutions to some standing objections to the existence of this moral conceptual entity. For instance, Strawson objects that

you will scarcely say that ethical intuitions are infallible; for ethical disagreements may survive the resolution of factual disagreements. . . . So your use of the language of "unanalysable predicates ascribed

in moral judgment to particular actions and states of affairs" leads to contradiction. For to call such a judgment "non-infallible" would be meaningless unless there were some way of checking it; of confirming or confuting it, by producing evidence for or against it. But I have just shown that your account of these judgments is incompatible with the possibility of producing evidence for or against them. So, if your account is true, these judgments are both corrigible and incorrigible; and this is absurd. (1949, p. 27)

OCU's version of Moorean moral objectivism does not assume that goodness is a simple, unanalyzable property, though this is consistent with available evidence; however, this version of moral objectivism does assume that either goodness or one of its parts is a simple, unanalyzable, ethical property. OCU's moral objectivism is therefore subject to a correspondingly modified variant of Strawson's objection.

This objection, even thus modified, is subject to a damaging trilemma. If the premise that ethical disagreements may survive the resolution of factual disagreements is taken to mean that ethical claims are not truth-functional, then the question is begged whether goodness is an objective property, claims of instantiation of which would either correspond with the world or

not, hence by the correspondence theory of truth be either true or false, respectively. Strawson's objection would thus be disqualified from use against OCU's moral objectivism.

If, on the other hand, the premise that ethical disagreements may survive the resolution of factual disagreements is taken to mean that ethical claims are not a posteriori propositions, then the moral objectivist of OCU can still hold that ethical claims are a priori propositions, disagreement about which is possible. As explained in ch. 2, ethical claims of the foundational sort in question though a priori may turn out to be either analytic or synthetic, depending on the metaphysical nature of determinables. While a priori propositions are noncontingent, humans are not universally able to discern their truth-values. Disagreement among reasonable people is common in this regard, particularly concerning the more complex of noncontingent propositions. Moreover, such disagreement can be reduced by consideration of evidence; conceptual analysis is a powerful evidence-gathering tool, as reflected in human progress in a priori fields such as mathematics and logic. There is no obvious reason why such disagreement should not arise regarding determinables. One may wonder whether a given entity (e.g., a transparent object) has the determinable property of coloredness, and one may wonder whether a given entity (e.g., a truth-telling) has the determinable property of goodness. Thus one could in this way call an objective moral judgment corrigible and non-infallible while agreeing that ethical claims are not factual

claims (a reply overlooked also by Hare - 1981, p. 69-70).

If, on the third hand, the premise that ethical disagreements may survive the resolution of factual disagreements is taken to mean that ethical claims are not entailed by non-ethical truths, then a subdilemma is generated.

If the premise that ethical claims are not entailed by non-ethical truths is taken in the sense merely that ethics and other subjects are separate subjects, then the moral objectivist of OCU can assent. No one expects truths about what is colored or smelly to directly resolve questions about what is noisy; nor should one expect truths about these properties to directly resolve questions about what is good. The various properties each ontically express a unique shared character aspect of the relevant range of objects. While propositions about each property have bearing on propositions about each other property to at least some extent via the vast web of belief, this influence may be subtle and attenuated; no direct relation of entailment need connect a given pair of properties. Even granting Strawson the dubious premise (see Bradley and Swartz 1979, p. 167) that there is evidence available for or against any non-infallible claim, one can simply point to the published evidence, for or against ethical claims, available within ethics via conceptual analysis. Part, though not necessarily all or even most, of this evidence consists of the coherence of a given ethical intuition with other related ethical intuitions, such as those of other people on the same topic. Strawson objects to

this line:

it is no good saying that, after all, the ethical judgments of other people (or your own at other times) may corroborate your own present judgment. They may agree with it: but their agreement strengthens the probability of your judgment only on the assumption that their moral intuitions tend on the whole to be correct. But the only possible evidence for the existence of a tendency to have correct intuitions is the correctness of actual intuitions. And it is precisely the correctness of actual intuitions for which we are seeking evidence, and failing to find it. (1949, p. 27)

Thus Strawson claims that the case for the correctness of an ethical intuition is no stronger than the evidence supplied from outside ethics. However, this claim is mere Cartesian skepticism applied to a particular human faculty, that by which we intuit ethical truths (and possibly other things, for we have found wanting the assumption that ethical intuitions require a *sui generis* conceptual mode). Cartesian skepticism is the view that we should doubt everything that can be doubted, i.e., that has not been proven from certain knowledge. But since no foundation of certain knowledge is available, everything must be doubted, including sensory or other epistemic intuitions, hence we have no

clue as to the nature of the world. The falsity of this conclusion suggests that we should take the argument as a reductio ad absurdum of its premise that we should doubt everything that can be doubted. Occam's Razor suggests the premise's contrary, that we should try to believe everything that our intuitions tell us to believe (see Rescher 1973, p. 54-64). The simplest explanation that fits the facts about why Chen seems to see a tree is that there is a tree and Chen is seeing it. On reflection we cannot accept all these potential beliefs, for some are mutually inconsistent; hence epistemology must winnow, amend, and supplement. This applies equally to the faculties by which we smell things, see things, feel things, and intuit things' moral qualities. Occam's Razor presumes against the existence of moral and all other intuitions, but once granted their existence it presumes in favor of their veridicality. This presumption is not forceless, either, for the existence of intuitions regarding the various non-ethical states of affairs generally is admitted (though with reservations, some warranted, e.g., Hare 1981, p. 9). Thus any given moral intuition is evidentially supported by itself, and perhaps also by cohering other moral intuitions and non-ethical evidence. The first two of these classes of intuition need not be forceless, as Strawson suggests they must; in fact, the first class is never forceless (see Baylis 1967, p. 447).

If, however, the premise that ethical claims are not entailed by non-ethical truths is taken in the sense (probably

the sense Strawson intended) that claims about goodness, e.g., that x is good, are not entailed by truths about goods, e.g., that x is a pleasant experience, then again the question is begged against OCU's moral objectivism, disqualifying Strawson's objection from use against it. While different properties need not stand in relations of entailment, they sometimes do; e.g., if x is square then x is rectangular, or, if x is a sound then x has a pitch. One should not assume without argument that different properties exactly one of which is ethically fundamental are unrelated by entailment (aside from the weak but everpresent presumption by ontic parsimony against entities in general, including entailments).

Since the premise expressed by the locution that ethical disagreements may survive factual disagreements has to be taken some way or other, yet none of the various ways discussed, each a rendering, from a representative set, of a philosophically popular way of taking such phrases, yields an interpretation of the objection that damages moral objectivism, the objection fails.

Strawson's objection, developed in the somewhat positivist, linguistically-oriented climate of mid-twentieth century philosophy, is complex and could be analyzed at length. However, it suffices for the moral objectivist of OCU to show as above that if ethical intuitions are taken to use as their intrinsically ethical parts mental items in a conceptual mode of a familiar non-imaging sort, the abstract mode, rather than as in

a sui generis conceptual mode unique to ethics, then Strawson's puzzles are seen to apply equally to goodness and to many familiar determinable properties, for many familiar determinables are conceivable only abstractly (e.g., coloredness and triangularity). While the existence of even non-ethical determinables is uncertain, their metaphysics and phenomenology are better-understood than those of ethical properties, allowing many people to get a satisfactory intuitive idea how Strawson's objection might be misplaced against determinables, and how it would hence be misplaced against OCU's moral objectivism. As with the property of goodness, the existence of the concept of goodness is on metaphysical grounds no less probable than the existence of many familiar determinable properties and their concepts. Since no one has provided a better alternative explanation of resemblance classes of determinate properties, no one has shown that a world without determinables is less mysterious than a world with determinables. OCU's moral objectivism, including the positing of a concept of goodness, is therefore a currently viable metaethic, available for use by utilitarianism and other comprehensive ethical theories.

It might be wondered why, if OCU does not hold that moral judgments are entertained in a sui generis conceptual mode, as intuitionism holds, I have bothered to address Strawson's objection to intuitionism. The reason is that I regard the distinction between intuitions of familiar mode and sui generis intuitions as a distinction of little practical use, just as I

regard the natural/non-natural distinction (and for that matter, the monism/dualism distinction - see §1.3). The precise conceptual mode by which any concept is entertained is in some respects unique and in other respects like other conceptual modes. Occam's Razor counts against modes or other items posited to be unique in too many respects, so as not to multiply determinables unnecessarily. However, vague impressions of just how unique an item is are almost as likely to reflect unconscious prejudices as to reflect anything of much epistemic weight. For instance, early estimates of humans' biological uniqueness led many 18th century naturalists to classify humans as the sole occupants of a taxonomic kingdom parallel to the kingdoms of animals and plants (see Gould 1993, p. 29), whereas modern opinion is that humans and other animals are not only members of the same kingdom, but also, in the case of, e.g., humans and chimpanzees, members of the same phylum, order, and family, differing only at the generic level (Szalay and Delson 1979). While vague impressions are better than nothing as evidence (i.e., are 'data' sensu coherentist epistemology - Rescher 1973, p. 57), in the present case we are not so desperate, for we have many analytical tools and much detailed empirical evidence at our disposal. We are thus better off spending less intellectual energy examining these vague impressions, and more energy examining the conceptual mode or modes actually or potentially used in ethical thinking, and using these particular results to ascertain the uniqueness of such modes, and whether they are

morally objective. This is not to suggest that we should be epistemic slaves to our traditional moral concepts or language, hence do other than "frame our questions clearly and then go out to find answers, letting the chips fall where they may", as Brandt suggests intuitionists mistakenly do other than (1979, p. 3); rather, it is to suggest that when answering the clearly-framed question how closely moral concepts, including objectivist concepts if any, resemble other concepts, we are better off to examine such concepts than to examine (with as much energy, for all data have some weight) our prior intuitions about how close this resemblance might be. Brandt's objection succeeds not against modern intuitionists of the epistemological coherentist persuasion (Sleigh 1984), but against traditional intuitionists such as Ross (1967, p. 269), who place unwarranted weight on human moral intuitions, largely due to their mistakenly thinking that intuitions from outside ethics can have no bearing on ethical matters (e.g., Ross 1967, p. 277). Positions midway between these extremes, e.g., Rawls (1971, p. 34-49), are deficient to the extent that they embrace this mistake. Many who are not intuitionists nonetheless treat traditional intuitionism as if it were the best or only version of intuitionism, hence join with Brandt in overlooking the strongest variant of the theory they dismiss (e.g., Hare 1981, p. 12, 40).

2. Entertaining concepts of determinables

A skeptic of OCU's moral objectivism might allow that determinables, whether any exist or not, have been supplied with a fairly comprehensible metaphysical structure, hence allow that a property of goodness construed as a determinable cannot be ruled out on metaphysical grounds. However, the skeptic might ask why, if this property exists and has, as do so many properties, a concept available for sentient entertainment, more humans have not claimed to have entertained such a concept:

If our inability to identify ethical qualities is the result of our looking for the wrong kind of thing, once this error is exposed and we stop making it, the difficulty should be overcome. But it is not. . . . Whether or not we are aware of non-natural qualities, we are certainly never cognizant of our awareness, and this must also be explained. (Sikora 1962, p. 63)

Nowadays few people, even, as will be seen, among the best moral philosophers, do report entertaining the concept of goodness. As with reports of UFO abductions, if a striking event is reported seldom enough, one may reasonably wonder whether these few reports are mistaken, allowing ontic parsimony to snip away another unnecessary entity.

Regarding the contemporary scarcity of reports of

entertainings of a concept of goodness, it should be noted that while nowadays Moorean moral objectivism is philosophically unpopular, only a few decades ago Moore's influence was great, and Moorean moral objectivism prospered, with reports of concepts of objective moral entities being common. Since humans are so easily influenced by popular theory on this matter, we should not pay undue attention to the concept's contemporary reportage rate when judging whether the concept exists. The exact epistemic significance of fluctuating popularity in an entity's concept may in principle be scientifically ascertained (Sleigh 1993, ch. 10); however, such research has not yet been completed respecting the concept of goodness.

Philosophers who aren't aware of the concept of such a property aren't expected to take it on faith that others are aware of it, they are expected to take it on evidence. For them, goodness and its concept are purely theoretical entities, such as are commonly posited in many disciplines. For no obvious good reason, purely theoretical entities have been posited frequently in, e.g., physics, with scarcely a murmur of dissent from philosophers, whereas extraordinarily heavy weather has been made of the same move in metaethics. This double standard should be abandoned, presumably by being to some degree less receptive to entity-positing in physics and to some degree more receptive to entity-positing in metaethics. OCU's moral objectivism is highly parsimonious, positing only goodness and its negative twin (or negative half), badness, and the usual attendant entities, such

as the concepts of goodness and badness. It is difficult to see how a metaethic could posit fewer entities yet still get the explanatory job done; a metaethic that posited no entities whatever would be as needful of scrutiny as a theory of physics that posited no atoms, molecules, or other microparticles whatever.

Regarding the rectification of some error about identification of the concept of goodness; previous purported rectifications may not have worked, yet if goodness is a determinable property then this element of a full appropriate rectification has not yet been tried widely. Now that this element has entered the philosophical literature (Sleigh 1992), changes are possible.

Since the quantity of reports of entertainings of the concept of goodness fluctuates beyond much current epistemic usefulness, let us turn to their quality. The majority of the best moral philosophers, from Aristotle to Mackie, do not report such a concept. However, some excellent thinkers do report entertaining concepts of objective moral properties. Most prominently, Moore writes that

Whenever (one) thinks of 'intrinsic value' or 'intrinsic worth', or says that a thing 'ought to exist', he has before his mind the unique object - the unique property of things - which I mean by 'good'.

Everybody is . . . aware of this notion,

although he may never become aware at all
that it is different from other notions of
which he is also aware. (1903, p. 17)

Kant is less clear in his endorsement of concepts of objective moral entities. He speaks as if value were an objective, ontologically independent component of the universe, readily conceived; he says that

a good will . . . would . . . shine like a
jewel for its own sake as something which
has its full value in itself. . . . this
Idea of the absolute value of a mere will
(1948, p. 62)

and that

it is impossible to conceive anything at
all in the world, or even out of it, which
can be taken as good without qualification,
except a good will. (1948, p. 61)

and he uses in an existence claim (the German equivalent of) the phrase "the concept of the good" (1948, p. 81). However, Kant also uses the phrase "the concept of duty" (1948, p. 65), but later cautions: "we leave it unsettled whether what we call duty may not be an empty concept" (1948, p. 88). Kant's position thus seems to be that there appears to be a concept of goodness, but that appearances can deceive in such matters (a position that agrees with my personal phenomenology; if it turned out that no concepts of intrinsic ethical entities exist and that no

conceptless entities exist, then I would retreat to a roughly Nagellian version of non-entifying moral objectivism, sufficient to ground hedonic utilitarianism). However, it is uncertain that Kant meant anything more by 'the concept of the good' than 'the correct conceptual analysis of goodness', which is consistent with a naturalistic or morally skeptical reduction of the concept of the good to nonmoral concepts, i.e., those other than of objective intrinsic worth.

The examples of Moore and his ilk suffice to show that whether or not a concept of goodness exists, a concept does exist that is either the concept of goodness or else a concept so similar to the concept of goodness as to be able to fool some experts. If the latter, then the place in our ontology for a concept of goodness is still vacant, hence the concept of goodness is in this sense an 'empty concept'. A mere handful of examples such as Moore suffice to show that a goodness-seeming concept exists, because if we assume, as seems true, that these persons were sincere, articulate, sane, intelligent, benevolent, well-educated reporters of their mental states, who had given sustained hard thought to moral matters including moral phenomenology, then only some exceedingly remotely causally possible error could neutralize the tendency of such introspective reports to be veridical. It is as likely that Julius Caesar came and conquered but did not see as that Moore wrote the words quoted above but did not entertain concepts which he examined carefully and took to be concepts of fundamental,

objective, intrinsic moral entities. Examples of a few such people multiply the probability of truth beyond reasonable doubt - these people may or may not actually have entertained objective moral concepts, but they definitely thought that they had.

It might be objected to this line that evidence of similar strength is available for the existence of concepts of, or concepts seeming to be concepts of, all sorts of dubious property, from intrinsic holiness to intrinsic vital force. The moral objectivist of OCU may respond by accepting the existences of all these other concepts. Accepting such concepts is far from accepting any associated properties or their genuine concepts. Most such concepts presumably will turn out not to be of the purported properties, so that the moral objectivist of OCU need accept only a few of the dubious properties - the few that turn out to exist.

Against this sort of response Strawson says that if some people don't associate any concepts with words such as 'good', but others do, then such words "have quite a different meaning for one set of people from the meaning which they have for another set. But neither of us believes this" (1949, p. 25). However, Strawson gives no reason for not believing that such words are being used ambiguously. In fact, contra Strawson this explanation is very plausible, for ambiguities of this sort are very common. Etymology shows that word meanings change greatly over time, and in many cases a mechanism of change is that

concepts are gained or lost from word meanings. Such linguistic change occurs gradually, often with more than one meaning in popular use at the same time. As words slowly bifurcate, competing meanings diverge, instantiating serially many degrees of similarity, until separation occurs, producing two words that may or may not be spelt the same or pronounced the same, but definitely don't mean the same. At a certain point in this process, a human linguistic group may find that some of its members are using one meaning for a word, some are using another, and neither set wants to relinquish the claim of unique appropriateness for its favored meaning for some key contexts. Such a position has been reached in English for the word 'line', by which some mean a straight edge and some mean a straight or curved edge; the truth or falsity of the parallel line postulate (that for a given point and a line not containing the point, one and only one line containing the point is parallel to the other, given line) depends on which meaning is used. In a disagreement involving a folk-theoretic word, often one party will take the word to mean the concept of a substantive folk-theoretic entity, and the other party will take it to mean nothing, for they think not only that no such entity exists or could exist, but also that no suitable coherent concept is available, not even a self-contradictory concept. For instance, the concept of a square circle is, though incoherent because self-contradictory, nonetheless composed of coherent elements; by contrast, the concept of a soul is, to me, a vacuous opacity - not an

incoherent composition of coherent elements, not even a blurry mess. I can grasp nothing of the alleged concept; the word 'soul' seems a mere place-holder in a metaphysical theory. To some people, e.g. Hare 1981, p. 83-86, it seems that 'good' as Moorean moral objectivists use it is just a vacuous place-holder in a metaethical theory. Thus such people and Moorean moral objectivists mean quite different things by 'good', hence 'good' is ambiguous. A like ambiguity seems to obtain of the folk-theoretic word 'mind', which most of us take to mean the concept of a familiar substantive entity, but which Ryle seems to take to mean, in this context, no concept, for he thinks mental states to be entities which not only don't and couldn't exist, but are vacuous "occult" entities (1949, ch. 1; p. 51). Strawson's thinking it implausible that such ambiguity could account for a moral case such as that of 'good' is perhaps attributable to the fact that both parties know how to use moral words such as 'good', and may use them almost identically in most contexts. To the extent that meaning is equated with use, as it often was in the philosophical climate prevailing when Strawson wrote, it will seem implausible that people could use a word similarly yet mean it quite differently.

It should be remembered that theories contain concepts, that there are many times more false theories than true theories, and that in many cases the false theories are false because some of their concepts are confused. These confused concepts are real and legion, and it should surprise no one if such a concept turns

up in ethics posing as the concept of goodness. Thus skeptics of Moorean moral objectivism misplace their efforts if they hold that Moorean moral objectivists typically don't really entertain any concept when they think that they are entertaining the concept of goodness. Instead, any such skeptics of Moorean moral objectivism should admit that Moorean moral objectivists usually do at such times entertain a concept seeming to be a token of the concept of goodness, and concentrate on proving that such entertained concepts are not tokens of the concept of goodness.

Regarding the question whether we are cognizant of our awareness of ethical properties: while OCU's moral objectivism entails nothing on this issue, some Moorean moral objectivists in fact claim to be cognizant of their entertainings of concepts of objective moral entities. However, this cognizance seems far from inevitable; it is curious how some avowed skeptics of Moorean moral objectivism turn out to advocate analyses of personal choices and their grounds that closely parallel the structures of Moorean moral objectivism, leading one to wonder whether these skeptics may not have entertained concepts of objective moral entities without being cognizant of the fact. Mackie, for instance, is famous for his moral skepticism, claiming

that there are no objective values, . . .
that no substantive moral conclusions or
serious constraints on moral views can be
derived from either the meanings of moral

terms or the logic of moral discourse.

(1977, p. 105)

Yet Mackie also holds that hypothetical imperatives, statements of the form 'If you want X you ought to do Y', are true; e.g.,

if someone wants to get to London by twelve o'clock, and the only available means of transport that will get him there is the ten-twenty train, and catching this train will not conflict with any equally strong desires or purposes that he has, then he ought to, indeed must, catch the ten-twenty.

(1977, p. 65-66)

Mackie is well aware of Hume's dictum (1888, p. 469-70, = Treatise, Book 3, Pt. 1, sec. 1) that one cannot derive an 'ought' from an 'is' because doing so would involve introducing some new relation joining agent and act, this new relation being deduced from others entirely different from it, which seems inconceivable. However, Mackie claims that in his derivation

no 'new relation' is involved. 'Ought' . . .

says that the agent has a reason for doing something, but his desires along with these causal relations constitute the reason.

(1977, p. 66)

If so, then Mackie is a moral objectivist by his own account. It is objective fact that an agent has the desires it has and that the relevant causal relations obtain; since these constitute,

apparently noncontingently, a reason for the agent to act, and since some but not all of an agent's available acts receive this objective endorsement, a substantive, serious constraint on agent conduct objectively exists. Since agent conduct and constraints thereon are the bulk of the subject of morality, Mackie has proposed what he calls an objective morality.

Let's distinguish the strongest general version of moral objectivism, which posits at least one objective, intrinsically ethical entity, from the less strong general version of moral objectivism, which does not, but nonetheless holds that facts can be objective reasons for action, so that some acts are recommended by rationality over others; call the former version 'entifying moral objectivism' and the latter version 'non-entifying moral objectivism'. Entifying and non-entifying moral objectivism are jointly exhaustive of ethical theories by which facts unequivocally demand certain choices of moral agents; call these ethical theories versions of 'strong moral objectivism'. However, there are proponents of weaker theories by which moral agents should rationally follow certain objective procedures in coming to moral decisions, though the outcome of these procedures is unspecified; call these theories versions of 'weak moral objectivism'. 'Moorean moral objectivism' is formally defined as the form of entifying moral objectivism that posits a Moorean property of objective intrinsic goodness, its negative twin (or negative half), badness, perhaps their concepts, and no other underived ethical entity. 'OCU's moral

objectivism' continues to refer to the version of Moorean moral objectivism that includes the metaethical detail argued for in this thesis. 'Moral objectivism' will refer to the union set of strong and weak moral objectivisms. Moral skeptics will be identified particularly, e.g., 'skeptics of entifying moral objectivism'.

Mackie's position is a form of non-entifying moral objectivism, not all that different from OCU's form of entifying moral objectivism; only three major differences separate them.

First, while Mackie holds that only a given agent's present desires generate reasons for that agent to act, OCU holds that any sentient's desire at any time may generate reasons for the given agent to act (even possible future desires, e.g., as when Jones would desire to eat this chocolate if I revealed it and gave it to Jones; on possibility and moral judgments, see Moore 1912, p. 27; Sikora 1978, p. 126). Mackie discusses part of this alternative:

Do the desires . . . of other people . . .
constitute a reason for me to do something,
if I can, or to try to do something to
satisfy those desires . . . ? . . . if we
recognize this as a further class of reasons,
independent of any desire that I now have to
help these other people, we are again
bringing in the requirements of something
like an institution: an established way of

thinking, a moral tradition. . . . However
. . . , the conclusion that is firmly
established will be only of the form: This
institution requires such and such an action.
If we move to a prescriptive interpretation,
we shall be speaking within the institution.
But nothing logically commits us to doing
so. . . . (1977, p. 78-79)

However, this defense fails since the notion of a metaphysically unitary 'I' being at once the subject of a present desire, the knower that act Y will bring about state of affairs X, the actor who might do Y, and the enjoyer who would be satisfied if X obtained, is equally an institution to which no one is logically committed. What Mackie identifies as a reason of prudence (a reason of no benevolence; of the purest self-interest) is in fact a relation between desires, cognitions, actions, and satisfactions that are ontologically independent and not obviously morally connected.

First, those elements of a Mackiean reason of prudence that are cotemporal are not thereby bound to each other morally, no more than one's present self is morally bound to one's future self. As Mackie says, regarding the latter case,

We can indeed say that . . . (someone) now
has a (prudential) reason for an action which
will tend to satisfy not any desire which he
now has, not even a present desire that his

future desires be fulfilled, but only a desire which he knows he will have later . . . and that (other things being equal) he ought to act. . . . But in saying this we are leaning on our concept of the identity of a person through time and the associated expectation that a human being will behave as a fairly coherent purposive unit over time. . . . Human beings are more likely to flourish if they show such purposive coherence over time, so that it is not surprising that we have this useful cluster of concepts and expectations. (1977, p. 78)

For similar evolutionary reasons, it is no more surprising that we have a useful cluster of concepts and expectations about the identity of a time-slice of co-accessed mental states. However, it is logically possible for some mental states to regard themselves as morally dissociated from other, co-accessed mental states (e.g., Nagel 1970, p. 75, 126), and apparently causally possible as well, for our mental hospitals yield apparent examples of such sets of awarenesses lacking coherent purposiveness. It is a mere useful social fiction that without exception we humans are psychologically cohesive persons. Even those minds which are psychologically cohesive thereby manifest only certain familiar universals such as resemblance between their composing elements; any inference beyond these familiar

universals, e.g., to relations of duty or reasons to act, needs explanation.

Thus Mackie cannot endorse what he identifies as reasons of prudence yet criticize those who think that the desires of others or of one's future self may help form reasons for one's action; there is no need for the latter views to assume the warrant of any institution, no more than Mackie need assume the warrant of any institution when positing objective prudential reasons for one of two co-accessible mental elements to will the fulfilment of the desire of the other. An institutional warrant would, as Mackie says, fail, but objective reasons exist anyway, on other, better warrant, between co-accessed mental elements and between my action and others' desires.

It might be objected that there's nothing odd about treating co-accessed mental elements as morally connected, for there's nothing imaginary about the co-accession, or selfhood, of the bundle of mental elements. However, the mental elements bundled together in one co-accessed set of awarenesses are often disparate and conflicting. Consider one who, having been raised in a prejudiced society, finds mentally accessible some negative desire toward a certain minority. This person might repudiate such a desire but be unable to eradicate it from consciousness. If Mackie is right, then such a person is stuck with the prejudiced desire, and out of self-interest must act to satisfy it. But why cannot such a person say 'that desire is no more part of me than a shrapnel fragment; it's in me but it's other,

so its satisfaction is no gain for me'?

Second, and less controversially, not all of the elements of a Mackiean reason of prudence are cotermporal. The desire and cognition take place before the action, which takes place before the satisfaction; hence Mackie's reasons of prudence also involve moral relations crossing temporal gaps. This second consideration is by itself sufficient to show that Mackie's reasons for action are the same as altruistic reasons for action in their status as objective, intrinsic, rational moral recommendations, and that Mackie's metaethic is a version of strong moral objectivism.

A determined skeptic of strong moral objectivism might accommodate this objection by renouncing all objective reasons for action, even those reasons seemingly generated by desires entertained here and now (i.e., reasons either justified by logically antecedent 'unmotivated' or 'basic' desires, or explanatorily presupposed by 'motivated' or 'nonbasic' desires - see Nagel 1970, p. 29; Foley 1978, p. 69; Gewirth 1978, §2.1; Dancy 1991, p. 415-16), thus rejecting even Mackiean reasons for action. However, strong moral objectivists such as Moore and Mackie have a deep contrary intuition that there are such things as objective reasons for action, even though such strong moral objectivists are divided on questions such as whose desires help form reasons for whose action. The conceptual relations posited in these reasons seem knowable a priori, hence to those intellectually equipped, self-evident; to some, perhaps Mackie,

the objectivity of such action-guiding considerations is so transparent that it is taken without question, without explanation, without cognizance of the awareness. It is easy for humans to incorrectly interpret their reasons for action (Nagel 1970, p. 82).

Another moral philosopher who is in practice an endorser of prudential reasons, hence a strong moral objectivist, is Rawls, who claims that:

each person must decide by rational reflection what constitutes his good, that is, the system of ends which it is rational for him to pursue. . . . (1971, p. 12),

that:

we have . . . substituted for an ethical judgment a judgment of rational prudence. (1971, p. 44),

and that:

it is an open question whether the principle of utility would be acknowledged. Offhand, it hardly seems likely. . . . Since each (person) desires to protect his interests, . . . no one has a reason to acquiesce in an enduring loss for himself in order to bring about a greater net balance of satisfaction. (1971, p. 14)

The second major difference between Mackie's moral

objectivism and OCU's is that while Mackie does not stress the importance of desire-satisfaction produced when the causal means to the desired state of affairs is effected, OCU's moral objectivism does: the reason the agent has for doing Y, which will bring about X, which is desired, is that X will (it is hoped) satisfy the desire. The unique importance of desire-satisfaction (i.e., pleasure - see §6.1) may be shown by thought-experiments. For instance, consider one who desires to meet a certain movie star; does a friend who knows that the movie star is obnoxious in person have reason to introduce the fan to the star if the opportunity arises? Such a meeting would produce the object of desire but not satisfaction; the desire would not be quenched appropriately. Mackie's line is that examining such variations is pointless since there are no objective moral requirements which might or might not be fulfilled (1977, p. 77); however, it is just as easy to hold that since examining such variations is not pointless, there are objective moral requirements which might or might not be fulfilled. Perhaps one has more reason to have love and lose than never to love at all, and surely one has more reason still to love and win.

The third major difference between Mackie's moral objectivism and OCU's is that while OCU holds that the objects grounding objective reasons resemble each other so as to form a natural kind, and interprets such resemblance classes as determinates whose resemblance finds ontic expression as a determinable property, Mackie holds his set of objects grounding

objective reasons to be merely a family resemblance set, hence with no determinable property instantiated by each member of the set. I have already given OCU's evidence for the existence of determinables, and the rest of this third difference is merely possible disagreement about whether desire-satisfactions form a unified resemblance class, like the colors, or a mere family resemblance class, like cats.

I have explained this parallel between Mackie's moral view and OCU in some detail in the hope that moral skeptics of all sorts may see how entifying moral objectivists see themselves, not as like theists in the age of Darwin and the atom, but as realists who, like many, see objective reasons for action in the world, and, like not so many, require theories to assign entities to any objective structure posited. It is not at all clear which metaethical view is true; what is clear is that the metaphysical issues on which OCU disagrees with Mackie and many other self-proclaimed 'moral skeptics' are comparatively minor, and philosophically controversial, not dead and decided; in fact, many such 'moral skeptics' are merely non-entifying strong moral objectivists. Therefore Mackie's and other purported refutations of strong moral objectivism fail.

These three differences between OCU's and Mackie's accounts of objective reasons for action also provide a framework for classifying such moral philosophies. For instance, Brentano, a staunch moral objectivist, agrees with Moore, regarding the first difference, that objective reasons for action may cross temporal

gaps:

one must consider not only oneself, but also one's family, the city, the state, every living thing upon the earth, and one must consider not only the immediate present but also the distant future. . . . To further the good throughout this great whole as far as possible - this is clearly the correct end in life, and all our actions should be centred around it. (Brentano 1969, p. 32)

Brentano also sides with Moore, regarding the second difference, in that he does not think pleasure the only good thing: "some . . . hold, in opposition to what experience makes evident to us, that pleasure is the only thing good in itself" (1969, p. 30). However, Brentano agrees with Mackie and Moore in thinking that pleasure, broadly construed, is good: "we prefer joy to sadness. . . . Were there beings who preferred things the other way round, we would take their attitudes to be perverse, and rightly so" (1969, p. 22). Brentano, like Mackie, parts with Moore and OCU regarding the third difference, the issue whether goodness is a property. Brentano agrees with Moore that all goods are alike in their goodness, that the range of determinate goods share their resemblance equally:

There is a common concept for those things that are . . . called "good". . . . The concept of what is good in itself is thus

univocal in the strict sense and not, as Aristotle taught, univocal only in an analogous sense. (1969, p. 75-76)

However, unlike Moore and OCU, Brentano shies away from positing a determinable property ontically expressing this resemblance: "the good is that which is worthy of love, that which can be loved with a love that is correct" (1969, p. 18). In thus identifying the good with good things, Brentano declines the chance, at the natural place in his argument, to assert, hence implicitly denies, that good things have goodness distinct from their good-making properties. Brentano, like Mackie, is therefore a non-entifying strong moral objectivist. As the identification of Mackie as a strong moral objectivist shows, these three issues are useful in classifying moral theories, and can reveal surprising alliances.

Such alliances need revealing, for contemporary moral philosophy is clouded by widespread ambivalence on the issues whether entifying moral objectivism is true and whether strong moral objectivism is true. Physics was similarly clouded when it was found necessary to talk about light as if it were both a wave phenomenon and a particulate phenomenon.

The revelation of these alliances amounts to a remodelling of the contemporary orthodox moral landscape. For instance, while Brandt has often been regarded as a moderate moral objectivist somewhere between the extreme moral objectivism of Moore and the extreme moral skepticism of Mackie, Mackie is

revealed to occupy a position somewhere between Moore and Brandt. Moore is an entifying strong moral objectivist, Mackie is a non-entifying strong moral objectivist, and Brandt is a weak moral objectivist. Brandt is a weak moral objectivist (at least in places) because though he gives objective procedures for processing one's desires before pronouncing the resulting set 'rational desires' and acting on them, he specifies nothing about the nature of these 'rational desires' or the acts or consequences they bring about (Brandt 1979).

While some philosophers who have been aware of concepts of objective ethical properties seem not to have been cognizant of their awareness, others who have been cognizant of their awareness seem either not to have fully grasped the nature of the concept, or not to have correctly identified the set of objectively intrinsically valuable things. The disagreements thus produced among such philosophers may suggest the following objection:

there has been considerable disagreement among philosophers who think that they have perceived objective moral qualities both as to what sorts of things possess them and as to what objective moral qualities are. For example, some of them have thought that we have a moral sense, while others would deny this. Some hold pluralistic views as to the good, and there is considerable disagreement

among them as to what sorts of things are intrinsically good and to how they could be ranked, while others hold monistic views. And there is the difference between deontologists like Ross and utilitarians like Moore. This suggests that most of these philosophers thought that they were perceiving moral qualities when they weren't. Thus the group of philosophers who may have perceived objective moral qualities would seem to be very small indeed. (R.I. Sikora 1992, personal communication)

However, these disagreements seem no different in character from those among researchers in perfectly respectable areas of inquiry, where the researchers in question, despite their disagreements of the above sorts, nevertheless have uncontroversially entertained the same concept of the entity in question. Consideration of some details of these disagreements in their ethical forms will be found in various sections of this thesis (locations indicated); however, to reply to the above objection it suffices merely to provide harmless examples of the disagreements in non-ethical forms. The harmlessness of these examples shows that these sorts of disagreement are capable of harmlessness, hence that an ethical concept is not proven deficient merely because such disagreements can arise about it. First, for instance, just as futurologists disagree on whether we

entertain certain beliefs about the future by a sense of precognition or by reasoning, consciously or subconsciously, from knowledge of the present, even though futurologists of both persuasions are entertaining the same beliefs about the future, entifying moral objectivists disagree on whether we entertain concepts of objective moral entities by a moral sense (see §3.3) or by some other method, even though entifying moral objectivists of both persuasions are entertaining the same concepts of objective moral entities. Second, just as hedonists disagree as to whether there is one sort of pleasant thing or a plurality of sorts of pleasant thing, even though hedonists of both persuasions are entertaining the same concept of pleasure, entifying moral objectivists disagree as to whether there is one sort of good thing or a plurality of sorts of good things (see §6.1), even though moral objectivists of both persuasions are entertaining the same concept of goodness. Third, just as proponents of the synthetic a priori disagree as to what sorts of proposition have the property of synthetic a prioricity underlying the concept of synthetic a prioricity they entertain, even though these proponents of the synthetic a priori are entertaining the same concept of synthetic a prioricity, pluralistic entifying moral objectivists disagree as to what sorts of things have the property of intrinsic goodness underlying the concept of goodness they entertain, even though these pluralistic entifying moral objectivists are entertaining the same concept of intrinsic goodness. Fourth, just as

pluralistic hedonists may disagree as to how pleasant things could be ranked in intensity of pleasantness, even though these pluralistic hedonists are entertaining the same concept of pleasure, pluralistic entifying moral objectivists may disagree as to how intrinsically good things could be ranked in intensity of goodness (see §6.1), even though these pluralistic entifying moral objectivists are entertaining the same concept of intrinsic goodness. Fifth, just as philosophers of mind disagree as to whether intelligence is a functional entity or an entity of qualia, even though these philosophers of mind, though perhaps not fully grasping the concept of intelligence, are grappling with the same concept of intelligence, entifying moral objectivists disagree as to whether intrinsic value is intrinsic goodness (value in things that are not acts) or intrinsic rightness (value in acts), even though these entifying moral objectivists, though perhaps not fully grasping the concept of intrinsic value, are grappling with the same concept of intrinsic value. This fifth sort of disagreement is perhaps reducible to the third sort.

Thus the view that these disagreements in moral contexts suggest that most of these philosophers thought that they were perceiving moral qualities when they weren't turns out to be mistaken. Thus the group of philosophers who may have perceived objective moral qualities would not, after all, seem to be very small indeed.

3. The feel of goodness and the sense of right and wrong

Finally, the skeptic of OCU's moral objectivism might ask how it feels, in a loose sense, to entertain the concept of goodness. One cannot answer such a question directly, for the same reason that one cannot directly explain to a blind person how red and yellow look, or explain to one who has not tasted them how chestnuts taste, or explain to one who has never been drunk how alcoholic intoxication feels. Goodness has a characteristic experiential 'flavor' which is unique and can only be hinted at to non-experiencers of the concept. To paraphrase Nagel (1974), there is a way of being an entertainer of goodness that is not identical to any other way of being, and is in some ways completely inaccessible to non-entertainers of goodness. However, let me do what I can to explain by analogy what it feels like to entertain the concept of goodness - or at least what entertaining the concept I take to be goodness feels like.

As is often the case when trying to conjure up, as vividly and tangibly as possible, a concept of a determinable, one may usefully start by reviewing the concepts of some of the relevant determinate properties. One then fixes on that aspect in which these concepts resemble each other, on the way in which these concepts have that aspect wholly and completely and equally, unlike family resemblances, which admit of degrees of resemblance. When I follow this procedure for goodness, it strikes me that certain determinate states of mind are all

wonderfully fulfilling and worthwhile in themselves, positively sublime whether cheery or tragic. This common aspect dazzles with peace and contentment, yet also engages with change and energy, no matter how little or great the intensity of goodness involved (likewise, hotness is present equally in the sensations of both warm and boiling water). I realize that to those who lack such a concept of goodness, these analogies may be of no more use than a winemaker's claim that their wine tastes confident yet unassuming, or an advertiser's assurance that their product smells fresh as springtime, or a human's claim that red looks like a trumpet sounds; however, it is the best I can do. Poets, artists, and musicians are often better able than philosophers to provoke elusive mental states in audiences.

This feel of the concept of goodness helps account for the fact that goodness and humans, like pleasure and humans, are such that necessarily humans who grasp the concept of goodness (or pleasure) have some positive disposition toward goodness (or pleasure), though not necessarily to the exclusion of a simultaneous negative disposition toward goodness (or pleasure) (in agreement with Nagel 1970, p. 109-11). The feel of goodness also helps account for one's ability to commend by attributing goodness despite the coincident descriptive function of such attributions: this conceptual content is naturally commendatory (Baylis 1967, p. 447, agrees). I use the term 'naturally commendatory' in much the same sense that Brandt, who uses the term 'rational' to denote those acts one would do if one's

decision process were fully and errorlessly molded by available information, believes that:

the term in that sense will naturally be
commendatory . . . (1979, p. 1),

there is no incompatibility between a statement
being true or false and with a descriptive
meaning, . . . and its having performative (e.g.
recommending) force just on account of that
meaning. (1979, p. 15)

Again, a theoretical structure often treated as prohibitively mysterious when asserted of objective moral goodness (e.g., Hare 1981, p. 71-73) resurfaces asserted of objective prudence as if perfectly straightforward. In truth, such structures are neither straightforward nor prohibitively mysterious.

Those curious what it feels like to conceive of goodness and derivative moral entities such as rightness might also wonder what the faculty is by which we instantiate these moral concepts in occurrent thoughts, e.g., in moral judgments, and what it feels like to have such a faculty. This faculty cannot yet be specified precisely; OCU assumes only that ethical intuitions use as their intrinsically ethical parts mental items in a general conceptual mode of a familiar non-imaging sort, the abstract mode, rather than as in a sui generis conceptual mode unique to ethics. However, humans as a matter of empirical psychological fact typically do have a sense (taken broadly) of moral right and

wrong, subsuming what is often called a sense of justice. This psychological fact is routinely assumed by normative ethicists, e.g., Rawls (1971, p. 46). Possibly it is an illusory sense, like a sense of witch-detection which humans in some societies claim to have, but which is aimed at no natural (i.e., objective) kind. Nonetheless, humans typically have a moral sense, in that they claim to be able to tell, as a 'basic action' (sensu Danto 1963 or Goldman 1970), whether many states of affairs are right or wrong, and whether some states of affairs are more or less right or wrong than others. This sense of right and wrong is not omnineffective, for there are many states of affairs that appear to be matters of ethics, but on which one's moral sense yields no clear verdict as to whether the state of affairs is right or wrong, merely a set of conflicting intuitions. Likewise, one may be sure that an object held in one's hand has a weight, but be unsure whether that weight is above or below a suggested amount. Even if the sense of right and wrong is aimed at ethical natural kinds, it is logically possible that it is a completely unreliable sense, like the sense of precognition some humans claim to have, which is aimed at real (future) events but whose accuracy is not substantiated by any significant empirical evidence. Possibly more than one kind of sense of right and wrong exists; possibly some people make moral judgments by consulting sensed ethical properties, and others make moral judgments by other criteria, in which case their sense of right and wrong may reduce to applications of some other sense or

senses. Though the claim must be qualified in these ways, humans uncontroversially have a sense of right and wrong. The cause of their having such a sense is evolutionary (Simpson 1969, p. 134).

It is sometimes objected to the moral sense theory that we should but don't know of any organ by which moral sensing is done. According to OCU the primary organ by which humans exercise this moral sense is the brain. Cognitive theory is progressing rapidly, and it may soon be possible to specify just which parts of the brain perform just which functions in the process of making moral judgments. However, our current lack of the full details of these functions does not entail that no moral sense exists or that we don't know which organ is the organ of moral sense. We don't yet know the full details of how the organs of any senses work. For instance, we have a sense of sight - but all we know of the details are that the eyes feed certain neural impulses to the brain, and that certain neuron structures and brain regions seem involved in the processing of these impulses and consequent production of visual qualia. We infer these details from facts about, e.g., eye structure, brain blood flow and brain-wave production in the various brain parts during episodes of seeing, and from details from artificial intelligence theory about how neurons might act as inferential matrices. We know even less about our sense of, e.g., addition, whereby we mathematically intuit the sums of numbers. But details similar to those available about our sense of sight are also available, via the same current technology, about additive

sensing and moral sensing. Supposing these tests were done, and we knew how the neural impulses fed from the ears to the brain, in one asked a moral question, and how certain patterns of blood flow and brain-wave production were subsequently provoked - would this count as knowing the identity of the moral sense organ? If so, then the moral sense theory is on firm anatomic ground; if not, then the senses of addition and sight are also open to the no-organ objection, which hence is not strong enough to be a serious worry to the proponent of a sense for which other good evidence is available.

Those humans whose sense of right and wrong does not allow them to directly consult objective ethical properties may wonder what it feels like to have such a capacity. It feels no different from any other capacity to perform a basic action. One wills the moral evaluation to occur, and it does. Unless one is a specialist, one has no more idea why the act of will is followed by the evaluation than one has why another kind of act of will is followed by one's leg moving. It just happens, as one experiences in so many aspects of life. Virtually all of the experiential novelty is in the moral concepts, e.g., of goodness, themselves.

This sparse theory of sensing right and wrong troubles some; Nowell-Smith objects that:

the intuitionist's reply to the question 'How do I know what I ought to do?' is . . . unenlightening. For it turns out to be: "You

know what you ought to do by intuiting the non-natural characteristic of obligatoriness that inheres in certain actions. But, my dear sir, do not be alarmed by this mysterious phrase. It is only another way of saying that you know that you ought to do those actions." We know what we ought to do by knowing what we ought to do. Opium sends us to sleep because it has a virtus dormitiva. (1954, p. 43)

However, since, unlike the traditional intuitionist, OCU holds that moral intuitions may have much in common with other sorts of intuition, OCU can and does hold that due to such a resemblance, this objection to basic acts of moral judgment applies equally to numerous common mental basic acts. How does one know whether a word on the page spells 'bib' or 'did'? One just looks, and the judgment occurs correctly and with so little delay that one might mistake it for the looking itself were it not for the reports of those with cognitive deficiencies preventing them from making such basic acts of visual judgment. They look, they see the letters on the page, but nothing else happens, like a paralytic unsuccessfully willing their inert limb to move. As with any basic act, *mutatis mutandis*, one intuits goodness by intuiting goodness; if one has the appropriate faculty, doing it is this easy; if not, little in the way of coaching or explanation is likely to help much. Since Nowell-Smith's objection fails

against common mental basic acts such as some visual judgments, because such acts exist despite the objection's applying to them, the objection fails generally, including against the mental basic act of intuiting goodness. What Nowell-Smith interprets as explanatory vacuity is actually mere ontological parsimony: don't posit entities where appearances aren't complicated enough to warrant it.

It might be objected to the existence of a sense of right and wrong directly consulting objective ethical properties that a great many people report having no such sense, whereas with other senses such as the senses of sight or hearing, most people have such a sense.

I reply, first, that if, as argued in §3.2, prudential reasons for action involve objective moral entities, then since very many people think that they can see that prudential reasons for action exist, very many people are, whether they are cognizant of it or not, aware of objective moral entities, hence have a sense giving them such awareness. A second consideration is that even if a large proportion of people lacked such a sense, this would not necessarily be unusual. Some senses are more specialized than others, and may be lacked frequently. For instance, some specialized aural capacities are often lacked by humans; e.g., the capacity to sense and contemplate certain subtle, complex musical rhythms. Such a lack may afflict a great many people; so might it be with the sense of objective right and wrong. An individual human typically has very many fewer atomic

sensory capacities than the full set ever recorded in one or more humans. For causal reasons, the sets humanly instantiated tend to form family resemblance clusters in sense-space; however, many other sets are causally possible. The potentially spotty distribution of human capacities is illustrated by idiot savants, who are infantile or vegetative in most respects, but highly skilled in isolated capacities, such as the ability to do certain mathematical operations, or play complicated pieces on a musical instrument.

It might be also objected to the existence of a sense of right and wrong directly consulting objective ethical properties that almost everybody having a sense such as sight or hearing is cognizant of it, whereas, if the first part of my reply to the above objection is correct, many people having an ethical sense are not cognizant of it.

I reply that while some senses, such as sight and hearing, are broad enough and important enough to human living that their absence is immediately noticable, other sensory capacities are narrow enough and unimportant enough that an unreflective human, as most are, may easily not notice their lack. For instance, if the capacity to sense and contemplate certain subtle, complex musical rhythms is lacked by a human, all that may be noticed is that the human doesn't like classical music or jazz (see Moore 1903, p. 191-92). A broad sense such as sight or hearing is, strictly speaking, not one sense but the sum of a great many narrow, specialized senses. If all of such a set are lacked, the

consequences are striking, but if only one or a few are lacking, no one may notice, not even the human having the lack, whether or not the lack is widespread. A human lacking the ability to sense a certain shade of green may not find out unless professionally tested, for since they can sense most shades of green, the lack leads to few or no mistakes in everyday life. Most people lack the ability to sense certain shades of extreme violet, but they do not notice the lack, for it's news to them that such shades exist. Depending on how narrow the sense of right and wrong is, and how closely related to other human senses it is, the lack of this sense may not be very conspicuous, even to the human with the lack.

Conversely, people having a narrow sense that others lack may not notice the sense in themselves or its lack in others. If people are told by enough others, and pleased enough by the suggestion, that the decisions they made and actions they took on the basis of information gained by the sense in question were actually made and taken on some other basis, they often believe it and suppress awareness of the real basis. For example, many people who are experienced and skilled at character analysis by means of subconscious processing of subtle verbal and behavioral clues are persuaded that they actually came to their beliefs about the subject's character via telepathy or tarot cards or somesuch. As in the case of character analysis, so with ethics: a spectrum of cognizances results. Some people are keenly aware of the moral sense in themselves or others, while other people

are oblivious to the same. Regarding the objection under present scrutiny (that almost everybody having a sense such as sight or hearing is cognizant of it, whereas arguably many people having an ethical sense are not cognizant of it), it suffices that there are some other specialized senses, even socially important ones, that many humans possess but are unaware of possessing.

It might also be objected to the existence of a sense of right and wrong directly consulting objective ethical properties that people reporting having no such sense are able to live normal, independent lives, whereas with other senses such as the senses of sight and hearing, people having no such sense are thereby limited in their pursuit of normal, independent living.

I reply that the discrepancy, if any, is much smaller than claimed. First, people lacking even broad senses such as sight and hearing often learn to compensate considerably by careful use of their other senses, and need help only in minor ways. Second, people lacking a sense of right and wrong live lives constrained and impoverished in some important ways not generally recognized as consequences of their moral sensory lack. Such humans may lack benevolence, which in many situations prevents them from sharing the fruits of cooperation. Human populations tending toward a Hobbesian state of nature are noticeably more violent and shortage-prone than societies with high levels of kindness and helpfulness. Or such humans may lack prudence, consequently neglecting the means to their own welfare, both immediate and distant. Because the losses incurred by moral sensory lack are

so often long-term consequences of present error, these losses are often not recognized as such consequences, and present error is not recognized as such. Blind humans run the risk of walking into walls; morally sense-impaired humans run the risk of being mugged on the way home from work, or being impoverished in their old age.

4. The open question argument

Those who are able to entertain the concept of goodness can use their ability to satisfy themselves that the property of goodness exists. The argument involved is a variant of what has become known, adapting Moore, as the 'open question argument'. Moore writes that:

ethics aims at discovering what are those other properties belonging to all things which are good. But far too many philosophers have thought that when they named those other properties they were actually defining good; that these properties, in fact, were not simply 'other,' but absolutely and entirely the same with goodness (1903, p. 10),

that:

the attempt to define good is chiefly due

to want of clearness as to the possible nature of definition. There are, in fact, only two serious alternatives to be considered, in order to establish the conclusion that 'good' does denote a simple and indefinable notion. It might denote a complex, as 'horse' does; or it might have no meaning at all. . . . (1) The hypothesis that disagreement about the meaning of good is disagreement with regard to the correct analysis of a given whole, may be most plainly seen to be incorrect by consideration of the fact that, whatever definition be offered, it may always be asked, with significance, of the complex so defined, whether it is itself good. . . . (1903, p. 15),

that:

(2) . . . the same consideration is sufficient to dismiss the hypothesis that 'good' has no meaning whatsoever. It is very natural to make the mistake of supposing that what is universally true is of such a nature that its negation would be self-contradictory. . . . And thus it is very easy to conclude that what seems to be

a universal ethical principle is in fact an identical proposition; that if, for example, whatever is called 'good' seems to be pleasant, the proposition 'Pleasure is the good' does not assert a connection between two different notions, but involves only one, that of pleasure, which is easily recognized as a distinct entity. But whosoever will attentively consider with himself what is actually before his mind when he asks the question 'Is pleasure (or whatever it may be) after all good?' can easily satisfy himself that he is not merely wondering whether pleasure is pleasant. And if he will try this experiment with each suggested definition in succession, he may become expert enough to recognize that in every case he has before his mind a unique object, with regard to the connection of which with any other object, a distinct question may be asked. (1903, p. 16),

and that:

if we recognize that, so far as the meaning of good goes, anything whatever may be good, we start with a much more

open mind. . . . when we think we have a definition, . . . we shall start with the conviction that good must mean so and so, and shall therefore be inclined either to misunderstand our opponent's arguments or to cut them short with the reply, 'This is not an open question: the very meaning of the word decides it; no one can think otherwise except through confusion.'

(1903, p. 20-21)

I view Moore's expression of the open question argument as susceptible of a plausible interpretation, as follows. The experiential feel and flavor of the concept of goodness yield two premises to anyone properly able to entertain the concept of goodness (and to otherwise think intelligently): first, the premise that a proposition of the form ' $x(Gx \leftrightarrow Nx)$ ', where G attributes goodness and N attributes some non-ethical property or (non-trivially) self-consistent complex of properties, cannot be analytic; and second, the premise that such a proposition concerns relations between two positive, contentful ways of being. The first premise follows from the novelty and uniqueness of the experiential feel and flavor of the concept of goodness. That such a concept could turn out to be constructed entirely of bits and pieces of other concepts with other experiential feels and flavors seems as likely as that the concept of redness should turn out to be composed entirely of concepts of light vibrations

or popsicle sticks. From the first premise it follows that goodness is not identical with any non-ethical property or complex of properties, hence that no such definition of goodness can succeed. Thus goodness cannot be reduced to some less controversial entity from elsewhere than the realm of objective intrinsic value, as has often been attempted. The second premise follows from the stubborn tangibility of the concept of goodness; if it's not there, why won't it go away, and where is its experiential flavor and feel coming from? From the second premise it follows that goodness is not a conceptual void, a theoretical placeholder, but a substantial thing in the world, a property. Since goodness is not nothing and is not something else, it must be itself - an independent property adding a new and important way of being to our ontology. This argument is unfortunately unavailable to those unable to entertain the concept of goodness, but is compelling to those who have this ability, hence have epistemic access to the two key premises.

It does not follow from these results that it is not analytic that goodness stands in the determinable-determinate relation with determinate good things. As shown in ch. 2, it is possible that determinates carry within themselves, as a part, the determinable in question. If so, then a proposition of the form ' $x(Ax \rightarrow Gx)$ ', where A attributes the determinate good in question, could not also be of the form ' $x(Nx \rightarrow Gx)$ ', for A has ethical content, its good part, whereas N does not.

Brandt criticizes Moore's version of the open question

argument on the ground that Moore has not attended sufficiently to the distinction between overt and covert synonymy (as Brandt terms them). Brandt writes that:

two property-referring expressions mean the same if and only if they satisfy the following two conditions: (1) For every actual thing or situation, the speaker must, if called upon to judge, be willing either to apply both expressions, to reject both, or must be in doubt about both. . . . (2) The first test is not quite adequate, for it would permit us to say that "unicorn" and "centaur" mean the same. In order to avoid this difficulty, we must consider the speech responses of a person to things or situations that are not actual (1959, p. 160),

that:

Some writers appear to feel that matters are much simpler, that all we have to do is ask ourselves about our intentions - what we intended someone else to learn from what we said. But how are we supposed to find what our intention was, except by noticing what we would have been content to say instead of what we did say - except, perhaps, in

specially simple cases? (1959, p. 162),

that:

It is useful to mark a difference between the case of two expressions satisfying our criteria and also obviously meaning the same, and the case of expressions satisfying our criteria but not obviously meaning the same. In the former case we shall say they . . . are overtly synonymous; in the latter case we may say they only . . . are covertly synonymous (1959, p. 163),

that:

Moore's most important suggestion was the . . . "open question" test or criterion for sameness of meaning. . . . Suppose you have two terms, P and Q, and it has been suggested that they mean the same. Now, Moore said, a way to test whether they really do mean the same is this. Compose a question of this form: "Is everything P also Q?" (or the reverse). Then ask yourself whether this question is "intelligible" or "significant" or whether you understand what it would mean to doubt that the answer is affirmative. If the

question is intelligible and significant, Moore said, then the terms do not mean the same, for if they did, the question would be no more intelligible than would be the question: "Is everything P also P?". . . . Moore thought no naturalistic definitions (of ethical terms) pass this test.

Certainly "desirable" and "is desired by somebody" do not, for "Is everything that is desired by somebody also desirable?" seems to be an intelligible and significant question (1959, p. 164),

and that:

What can Moore mean . . . when he says that, in doubtful cases, we are to ask ourselves whether his question is intelligible, an "open" question? Perhaps what he means is simply this: that P and Q do not mean the same for a person if, when he asks himself "Is everything P also Q?" he is doubtful of the answer, or the answer does not seem to him obviously affirmative. . . . But if this is what Moore meant, then it appears he was mistaken at least for covert synonymy, for it might well be the case that two terms do covertly mean the same, as a given person

uses them, so that he could never correctly deny that everything that is P is also Q, but, if the person were not clearly aware of his use of terms, he might remain doubtful, and the answer might seem to him not obviously affirmative. . . . Furthermore, two terms might pass Moore's test even if they didn't mean the same. (1959, p. 165)

Brandt succeeds in showing that so far as the open question test is construed as a quick test of sameness of meaning for property-referring terms, it is deficient in the respects he mentions. It is unclear whether Moore's argument is seriously damaged by this demonstration, for it might be held that since Moore held that one must, to see that goodness is an indefinable property, by repetition become "expert" (1903, p. 16) in the application of the open question test to goodness and its various purported definitions, it is reasonable to explicate Moore's demand as including that one must apply Brandt's clarified test when arguing for synonymy, hence would not make mistakes by claiming synonymy quickly or unwarily. However, Moore is guilty at least of omitting responses to some reasonable if ultimately undamaging objections.

The merit of Moore's position is, however, not my present concern; my present concern is whether Brandt's objections apply to the construal of the open question argument I claimed above to show that goodness is an independent property adding a new and

important way of being to our ontology. That argument has two parts.

The first part uses the premise that a proposition of the form ' $x(Gx \leftrightarrow Nx)$ ', where G attributes goodness and N attributes some non-ethical property or self-consistent complex of non-ethical properties, cannot be analytic. This premise was claimed to be epistemically accessible to any intelligent person able to entertain the concept of goodness, and to yield the conclusion that goodness cannot be reduced by definition to some less controversial non-ethical property, simple or complex. Brandt's correct observation that detecting synonymy is seldom or never a simple process, available to the quick or unwary, does not damage this first part of my argument, for I argue not for synonymy, but for lack of synonymy. As Brandt remarks,

the proper procedure for showing that a definition is mistaken is to pose a "counterexample" - an example, actual or possible, to which all or most people would apply the one term but not the other. Clear counterexamples definitely prove that a definition is mistaken. . . . it is much easier to disprove a definition than it is to establish that one is correct, for a disproof (except where there are several senses) requires only a single counterexample. (1959, p. 162)

The lack of synonymy I argue for is between the term 'good' and some non-ethical term such as 'highly evolved'. I imagine a state of affairs including a highly evolved object such as a birch tree, but excluding any manifestation of the feel and flavor forced upon my consciousness when I contemplate the aspect in which various determinate goods fully and equally resemble each other. To this state of affairs I would be willing to apply the term 'highly evolved', but unwilling to apply the term 'good'. One does not become certain of such lack of synonymy overnight, but the longer I think about it, the surer I become. After consideration and rejection of a number of possible synonymies between 'good' and various non-ethical concoctions, one becomes aware of a sameness to the gaps of meaning, and aware of a growing conviction that the whole project is misconceived. Like one who considers and in turn rejects various physicalist definitions of a mental term such as 'red', one comes to see that two quite different sorts of concepts are being compared, and that consideration of further definitions of the sort in question is a waste of time.

The second part of my argument uses the premise that a proposition of the form ' $x(Gx \leftrightarrow Nx)$ ' concerns relations between two positive, contentful ways of being. This premise was claimed to be epistemically accessible to any intelligent person able to entertain the concept of goodness, and to yield the conclusion that goodness is not a conceptual void, a theoretical placeholder, but a substantial thing in the world, a property.

Brandt's correct observation that detecting synonymy is seldom or never a simple process, available to the quick or unwary, does not damage this second part of my argument, for I argue not for synonymy, i.e., that the two terms have the same meaning, but rather that each term has a meaning. Brandt does not specifically address the question how one is to tell whether a term has a meaning, but from his recommendation of the strategy of counterexamples for proving lack of synonymy, it may be extrapolated that a term has a meaning if a set of states of affairs can be specified to which all or most people would exclusively apply the term. While 'good' perhaps does not satisfy this criterion, it does satisfy a criterion similar but restricted to those who entertain the concept of goodness. One might even argue that 'good' also satisfies Brandt's unamended extrapolated criterion, on the ground that 'good' is used in several senses, and that the relevant sense is used appropriately by most or all of those who use it at all.

To this line it might be objected that it is not the case that most or all of those who entertain the concept of goodness apply the term 'good' exclusively to a certain set of states of affairs. The marked disagreement among certain Moorean moral objectivists about which things are good might be held to show that such people apply the term to rather different sets of states of affairs. However, this objection mistakes disagreement about which determinate things have a given determinable for disagreement about which things are the determinable in question.

People can disagree about whether transparency is a color without significantly disagreeing about what coloredness is; and they can disagree about whether knowledge is good without significantly disagreeing about what goodness is. When Moorean moral objectivists entertain the concept of goodness, they recognize that they have done so - whether or not they have entertained this concept in an appropriate context. The enduring popularity of hedonistic utilitarianism among Moorean moral objectivists evidences that they frequently, perhaps even usually, entertain the concept of goodness in appropriate contexts, hence are as warranted in their positing instances of a property of goodness as others are in positing instances of a property of coloredness when they see reds and greens. Their concepts are entertained at the right time, for the underlying property is also present.

Thus whether or not Moore's version of the open question argument is sound, a version supporting OCU is sound.

Thus in this chapter a satisfactory positive account of the concept of goodness has been given; hence the second objection to OCU is refuted.

Chapter Four

COMPONENTS OF UTILITARIANISM

1. Compatibilist utilitarianism

The third objection to OCU (and utilitarianism in general) is that the most plausible and popular utilitarian doctrines are forms of either act or rule utilitarianism, but that each of these doctrines suffers a repellent objection. AU is the doctrine that an act is right if it will result in at least as much utility as any other available act; RU is the doctrine that an act is right if endorsed by a rule from a set whose adoption by society will result in at least as much utility as the adoption of any other set of rules.

Some ethicists (e.g., Brandt 1963, p. 109-110; Harrod 1936, p. 147) are repelled by AU for its apparent implication that the *prima facie* intuitively justified rules of commonsense morality (e.g., that promises should be kept, that truth should be told, and that innocent persons should not be punished) deserve no special favor and on any given occasion should only be followed if doing so is likely to produce more utility than not doing so. Cynicism and social disorder seem more likely than harmonious happiness to result from taking these rules so lightly. Call

this the counterintuitivity objection. The intuitions violated can be construed in either of two ways; they can be construed as scientific intuitions that subscription to the rules in question must be beneficial, contra the claims of AU, or as moral intuitions that, beneficial or not, such rules must be obeyed. I shall discuss both construals of the relevant intuitions.

On the other hand, some ethicists (e.g., J.J.C. Smart 1973, p. 10) are repelled by RU for its apparent implication that rules should be followed even on occasions when doing so would not gain any utility. A moral code that makes demands unsupported by the only intrinsic value it (overtly) posits is unconvincing. This is the rule worship objection.

Thus both the chief forms of utilitarianism are seriously flawed; hence utilitarianism including OCU should be abandoned altogether.

In this chapter I defend a form of utilitarianism, the form assumed by OCU, that combines the main philosophical virtues of AU and RU while retaining the vices of neither.

Many ethicists are attracted to hedonic utilitarianism for its metaphysical parsimony and its positing pleasure, under some construal, as intrinsically good, and pain, complementarily construed, as intrinsically bad. For purposes of this chapter, intrinsic values may be interpreted either cognitively, e.g., as in OCU's moral objectivism, or non-cognitively, e.g., J.J.C. Smart 1973, p. 4. Both AU and RU, the two most well-expounded general forms of utilitarianism, are susceptible of hedonic

interpretations, non-hedonic interpretations, and of interpretations classing both hedonic and non-hedonic entities as utility. However, the interpretations of the third, hybrid sort rightly attract less contemporary support than purely hedonic interpretations, and interpretations of the second, purely non-hedonic type rightly attract almost no contemporary support, and hedonism in general is popular and plausible (see Goldstein 1989), so I shall confine the scope of my remarks to purely hedonic interpretations despite these remarks' mostly also applying to any partly or wholly non-hedonic form of utilitarianism.

To see that there is a form of utilitarianism more plausible than either AU or RU as standardly interpreted, we must now qualify the condensed definitions given above of AU and RU by distinguishing ontic from epistemic criteria. The ontic criterion for X-hood says what an X is. An epistemic criterion for X-hood says how an X may be known from other objects. For example, the ontic criterion for being a platypus is having each of very many disjunctive anatomic, behavioral, mental, and other properties (enough to completely describe platypuses). One epistemic criterion for being a platypus is being an egg-laying, web-footed mammal; another is being of the same species as this animal here (pointing at a platypus). For a given X there is only one ontic criterion, but there may be many epistemic criteria. Different epistemic criteria will most efficiently serve different observers. An ontic criterion may also be an

epistemic criterion for its object, but this need not be so for a given observer. For example, a blind human can tell a red tomato from a green tomato, but not by color.

The distinction between ontic and epistemic criteria has been called, respecting intrinsic goods, the distinction between 'conferring properties' and 'identifying properties' (see Baylis 1967, p. 446). I use the wider version of the distinction because it strengthens OCU metaethically to use premises of general epistemology and metaphysics rather than premises peculiar to ethics, for this makes it more obvious that no hidden, question-begging values are being smuggled into OCU's premise set. Ethical value is goodness and badness, and nothing else.

Hare makes a similar distinction between 'critical moral principles' and 'prima facie moral principles', these principles being criteria of right acts expressed in prescriptive form (1981, p. 38-41). Hare defines prima facie moral principles as relatively simple, general moral principles expressing prima facie duties and tailored to our human reactive attitudes so as to be useful at the intuitive level at which most practical moral decision-making is conducted. Hare defines critical moral principles as moral principles of unlimited specificity chosen under the constraints imposed by the logical properties of the moral concepts and by the non-moral facts, and by nothing else. While Hare's distinction works well within his system, and captures some interesting differences between the two classes of

moral principles we want distinguished, it is imperfect because it mixes moral and non-moral features in a confusing way. It mixes Hare's theory of how to choose moral ends with the scientific theory of how to promote ends once chosen. OCU is a form of strong moral objectivism, whereas Hare is (overtly) at best a weak moral objectivist, and perhaps not a moral objectivist at all; OCU and Hare thus advocate choosing moral ends in quite different ways. But given ends, OCU and Hare agree almost completely on the existence of two sorts of criteria for recognizing these ends, these criteria having separate places and uses in practical thought.

This distinction between ontic and epistemic criteria is crucial to utilitarianism. Any definition of utilitarianism including a criterion for a right act is ambiguous between one definition interpreting the criterion as ontic and another interpreting it as epistemic. AU and RU as defined above are ambiguous in this way. Thus we may distinguish ontic AU from epistemic AU, and ontic RU from epistemic RU. Epistemic versions must be indexed to particular kinds of moral agent to have truth values; e.g., 'epistemic AU, indexed to humans, is true' = 'for humans, epistemic AU is a maximally efficient epistemic criterion of right action'. Ontic AU and epistemic RU are true and the other two theories false; this view is compatibilist utilitarianism (CU). Ontic AU has been expressed in the utilitarian intuition that whether one is considering acts, sets of rules, or whatever, the bottom line in ethics is quantity of

benefit. Epistemic RU has been expressed in the utilitarian intuition that any moral code worth using as an everyday, practical guide must be both learnable and detailed, hence consisting of neither a huge set of rules nor only a single rule. Ontic AU and epistemic RU are not competitors but the two mutually consistent halves of a full utilitarianism. To be practical a morality must have efficient epistemic criteria; to be maximally efficient an epistemic criterion must pick out objects iff they satisfy the morality's ontic criterion; hence a morality must also have an ontic criterion (see Hare 1981, p. 46). Thus a full utilitarianism must have both epistemic criteria and an ontic criterion; epistemic RU and ontic AU supply these criteria. It is ontic RU and epistemic AU which are false and have done much mischief in ethics.

Ontic RU is liable to the rule worship objection. If following RU rules was itself the goal of ethics, then it would not need to be justified by its effect on utility or anything else, and could not be so justified because any justification would be via another ontic criterion, and there can be only one ontic criterion per object. Yet moral respect for rule-following for its own sake is mere worship. Some hold pleasure or happiness to be intrinsically valuable, hence the most precious objects in the universe; some hold knowledge, or love, or beauty, or honor to share this first rank; very few hold rule-following to be have anything but extrinsic value. Even Kant only considered rule-following valuable when it allowed the good will

to choose out of duty; the following of nonmoral rules he considered intrinsically worthless.

It might be objected to this criticism of ontic RU that since the appeal to utility is built right into the description of right acts, appeal is not being made to anything outside of right acts, such as their effects, hence that no second ontic criterion is implicitly appealed to to justify the first. However, if attribution of intrinsic value is thus built right into the description of right acts, then it must apply to rule-following as well as utility, for they are equally objects described as independent components of right acts; hence rule-following is still being valued for its own sake.

Epistemic AU is, when indexed to humans, liable to the counterintuitivity objection, for it leads us to lie, steal, break promises, punish innocent persons, etc. when our *prima facie* intuitions correctly tell us that we should not. The intuitions violated can, as mentioned earlier, be construed in either of two ways; they can be construed as scientific intuitions that subscription to the rules in question must be beneficial, contra the claims of AU, or as moral intuitions that, beneficial or not, such rules must be obeyed.

The bulk of this chapter addresses the first construal; I argue that AU taken correctly, i.e., ontically, entails that such rules should be part of our moral code because subscription to them would be beneficial. Thus AU does not claim that such rules should not be subscribed to.

Regarding the second construal of the intuitions in question, I argued (chiefly in ch. 3) that moral intuitions should be treated like any other sort of epistemic intuition. Epistemic intuitions are fallible, especially *prima facie* intuitions, but should not be rejected without reason; intuitions are the only raw materials we have to build a picture of reality from, hence carry a presumption of truth which can only be nullified by contrary epistemic data (see Rescher 1973, p. 55-59). In the present case, other relevant data turn out to corroborate our moral intuitions on the whole, including the intuitions that the rules of commonsense morality in question are justified. However, this corroboration comes at the price of a crucial qualification to the intuitions under discussion. It is false that the rules must be obeyed whether subscription to them would be beneficial or not. Rather, they must be obeyed because it turns out that subscription to them would be beneficial. The intuitions must be reinterpreted as intuitions that subscription to the rules would be beneficial. *Prima facie* intuitions, moral and otherwise, often need reinterpretation to be accepted, so this move is not *ad hoc*. However, some moral intuitionists will find this price too high; with them I must simply disagree, and hold that the unqualified intuitions are sufficiently lacking in coherence with our other epistemic intuitions that it is rational to reject them. Adherents of non-cognitivist forms of utilitarianism will have, if they support the second construal of the counterintuitivity objection,

to explain why moral intuitions count as data at all, but that does not concern cognitivist forms such as OCU.

Therefore epistemic AU is, when indexed to humans, liable to the counterintuitivity objection. The one-rule everyday moral code of epistemic AU might suit a society of near-omniscient beings (e.g., Hare's archangels - 1981, p. 44) who could perform the required calculations, but is ill-suited to beings of human intelligence. To complete this refutation of the counterintuitivity objection, let me explain the CU moral code in more detail to show why subscription to the commonsense moral rules in question would be beneficial.

The CU moral code consists of the single high-level rule 'maximize utility', explaining why the other rules should be followed, and a short list of lower-level, derivative rules for everyday practical use, largely consisting of rules of commonsense morality, such as 'keep your promises', 'tell the truth', and 'don't punish innocent persons'. The top rule 'maximize utility' is not for everyday direct practical use, for the potential for abuse and misunderstanding is too high. However, this rule is put to everyday indirect practical use, and to occasional direct practical use. The top rule 'maximize utility' is properly used chiefly: first, to help interpret lower-level rules in novel situations falling under their jurisdiction; second, to help settle new disputes between conflicting lower-level rules; third, to help design new rules to govern new situations as culture changes; and fourth, to help

decide when to abandon old rules made obsolete by cultural change. Uses of the first sort are the least significant and most common, for they involve no change in the code's rules, yet are encountered proportionately to the rate of cultural change, which tends to be high in a complex, civilized society. Uses of the second sort are more significant, and fairly common since novel conflicts between lower-level moral rules also arise regularly due to cultural evolution. Uses of the third and fourth sorts are the most significant and least common; citizens as democratic self-rulers should be encouraged to think regularly about the composition of their moral code, but since the list of moral rules for learnability is short, the rules will be based on extremely general features of human psychology, unlikely to change much in one human's lifespan, hence there will be little turnover of rules required. Contemporary examples of issues involving these four sorts of use of the top rule are, respectively: first, the question whether shortening a comatose human's life harms the human; second, the question whether the psychiatrist's promise to respect their patient's desire for confidentiality outweighs the psychiatrist's duty to tell the truth to the police about serious criminal activity learned of; third, the question how to formulate new rules about waste disposal now that so much waste is toxic to other species (and our own); and fourth, the question whether to drop the still-popular rule against marijuana use as an irrational taboo (see Boyd 1991).

The mystery how a human's sincerely trying to maximize utility at every turn could generate less utility than trying to do something else is explained as follows, paraphrasing Sikora (unpublished). When epistemic RU humans believe that it is wrong to break a rule, they in some choices believe that following the rule would maximize utility and in others not. When they believe that it would, they behave exactly as if they were epistemic AU humans. Epistemic RU humans only behave differently from epistemic AU humans when they follow a rule even though they believe that breaking the rule would have better results. But epistemic RU humans maintain that their overall pattern of behavior in following the rules will have better results than if they followed the epistemic AU single-rule moral code. This superiority can't be explained in terms of the part of their behavior that matches that of epistemic AU humans. The only possible explanation is that in the choices when they behave differently from epistemic AU humans by following a rule even though they believe that breaking it would have better results, they are wrong in their belief sufficiently often or their errors are sufficiently serious to make their policy of following the rules yield better results.

In practice epistemic RU humans will make numerous decisions to follow the rules when they believe that it would be better to break them, and sometimes they will be right that it would be better to break them. But if the overall importance of their erroneous judgments doesn't outweigh the overall importance of

their correct judgments they would have done better as epistemic AU humans.

Sikora calls this explanation the 'error theory'. An epistemic RU short moral code is more beneficial than the epistemic AU single-rule moral code iff the error theory is true. But, as Brandt argues, varied life experience of moral events strongly suggests that such a short moral code IS more beneficial. A human society that subscribed to it would be far happier than one that subscribed to the epistemic AU single-rule moral code because the latter leaves too much room for rationalization and because behavior in such a society wouldn't be sufficiently predictable. Thus the empirical evidence suggests that the error theory is not merely logically possible but actually true.

The phenomenology of refraining from doing what one thinks would maximize utility because one doubts that doing that act would maximize utility is odd; it is like the phenomenology of spearing fish - don't throw the spear where you see that the fish is, because that's where you doubt that the fish is (due to the bending of light as it passes through an air-water boundary at other than a right angle). The seeming paradox of thinking both that p and that $\neg p$ is resolved; one asserts and negates the same proposition, but not in the same thought even though sometimes simultaneously. One expresses the proposition via two separate mental channels, one asserting it and the other negating it. It might be objected that one should not accept apparent

fish-seeings, etc. as beliefs until they have been studied, but humans often do, treating them as fact for purposes of action, and not as a gamble but with confidence in the beliefs' veridicality. History shows that humans can even believe contradictories.

The strategy of not doing x in order to do x crops up often in human affairs (Tocqueville 1840, p. 148; Nagel 1970, p. 132). Utilitarians using an epistemic RU short moral code should refrain from breaking the rules even when it strikes them that doing so will gain utility for the same strategic reason that investors, whose motive is to gain money, should avoid investing in double-your-money-in-a-week schemes even when it strikes them that the scheme will work and gain them money. Some types of appearance are known to be less reliable than others. When two types of appearance regularly conflict, one must know which type to bet on, especially when these appearances are used as epistemic criteria in moral codes.

It is inattention to this and other matters of psychological strategy that misleads moralists such as Ross, who writes that:

utilitarians say that when a promise ought to be kept it is because the total good to be produced by keeping it is greater than the total good to be produced by breaking it. . . . Now, we may ask whether that is really the way we think about promises?
(1967, p. 275-76)

It seems not to occur to Ross that we may sincerely think about promises and the like in more than one way, according to the situation, yet without any deep inconsistency. This confusion remains widespread (e.g., Taylor 1989, p. 226), though it has been pointed out by utilitarians for some time (e.g., Sidgwick 1874, p. 413; J.J.C. Smart 1956, p. 346; and Hare 1981, ch. 2. Hare, p. 25, attributes the point to Plato and Aristotle).

Epistemic RU is logically independent of ontic AU, hence does not require CU. Every epistemic criterion must be justified by some ontic criterion, but there can be rational disagreement about which ontic criterion a given epistemic criterion is justified by. Rule utilitarians such as Brandt typically and plausibly ultimately appeal to utility maximization when justifying their short list of everyday rules, which suggests that epistemic RU is justified by ontic AU, but perhaps some other ontic criterion is more plausible. However, the existence of even one plausible candidate ontic criterion to justify epistemic RU refutes the rule worship objection. Given the lack of any other plausible candidate ontic criterion and the high plausibility of ontic AU (immune to the counterintuitivity objection to epistemic AU), CU is the best comprehensive utilitarian theory. OCU is a cognitivist version of CU.

2. Utilitarian rules

In this section I explain more precisely how different kinds of rule function in CU; in light of these functions, it will be seen that the currently prevailing forms of AU and RU, Smart's and Brandt's versions respectively, differ very little and are essentially the same theory. This demonstration will corroborate my claim that a form of utilitarianism, the form assumed by OCU, can be defended that combines the main philosophical virtues of AU and RU while retaining the vices of neither.

As suggested in ch. 1, CU has deep intellectual roots. The distinction between the one ontic and the various epistemic criteria for right action, and their compatibility within utilitarianism, were appreciated by Mill: "Whatever we adopt as the fundamental principle of morality, we require subordinate principles to apply it by" (1861, p. 32); by Sidgwick: "it is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim" (1874, p. 413); by Moore (1912, p. 30-35); and by Hare (1981, p. 43). Regarding recent utilitarianism, CU's reconciliation of AU and rules does not differ greatly from J.J.C. Smart's (1973, p. 42-57). CU differs from J.J.C. Smart's version (aside from the latter's overt incompatibilism) chiefly in holding that higher-level rules should be used more often in situations permitting careful reflection than J.J.C. Smart suggests, and in holding that higher-level rules should be used as more than mere rules of thumb, in that they should be invested with dignity and broken only in rare circumstances, chiefly emergencies, not

anytime prima facie analysis suggests utility would be gained.

J.J.C. Smart holds that

The act-utilitarian will . . . regard these rules as mere rules of thumb, and will use them only as rough guides. . . . He acts in accordance with rules . . . when there is no time to think. . . . When he has time to think what to do, then there is a question of deliberation or choice, and it is precisely for such situations that the utilitarian criterion is intended. (1973, p. 42-43)

However, as will be seen, there are sound game-theoretic reasons for taking some kinds of rule more weightily.

CU's reconciliation of RU and the ultimacy of utility maximization does not differ greatly from Brandt's (aside from the latter's overt incompatibilism). Brandt agrees that the bottom line in ethics is utility maximization:

if it is really true that doing a certain thing will have the very best consequences in the long run, everything considered, of all the things I can do, then there is nothing better I can do than this. . . . succeeding in producing the best consequences is a kind of success which cannot be improved upon. (1963, p. 121)

Reconciling this bottom line with the need for some rules compliance with which calls for utility to appear occasionally to be forgone is difficult; CU differs from Brandtian RU chiefly in holding that the rule 'maximize utility' would surely play a central role in the set of rules optimal for practical use by sentients of human intelligence. This role, herein called that of 'top rule', is essentially the role Brandt calls that of

a consistent and plausible "remainder-rule,"
that is, a top-level rule giving adequate
directions for all cases for which the
lower-level rules do not prescribe
definitely enough or for which their
prescriptions are conflicting. (1963, p.

133)

A top rule is an epistemic criterion distinguished by its sole occupation of the highest hierarchical level in a moral code's hierarchy of rules for practical use, and may or may not also be the ontic criterion behind the moral code's epistemic criteria. Brandt thinks it an open question which rule should be the top rule (remainder-rule) in a human society, whereas CU's top rule in such circumstances is, for reasons of human psychology, always 'maximize utility'.

These differences are minor; CU's main innovation is showing how these two theories, J.J.C. Smart's and Brandt's, are in all fundamentals consistent, and are essentially the same theory.

To see this similarity, let's consider in more detail the

role according to CU of the rule 'maximize utility'. As utilitarianism's ontic criterion of right action, the rule 'maximize utility' guarantees success if followed. However, as an epistemic criterion this rule is of limited value, in that in many contexts right acts can be identified more reliably by their contingent possession of some other property, e.g., that of being the keeping of a promise, which thus specifies a rival utilitarian epistemic criterion of right action. This structure is game-theoretically identical to that by which chess's ontic criterion of right action, the rule 'trap opponent's king', guarantees success if followed, but is of such limited epistemic value that the acts it specifies are usually more easily identified by means of other epistemic criteria, stated in rule form as, e.g., 'castle early in the game', 'advance weak pieces before strong pieces', and 'control the center of the board' (Hare 1981, p. 37-38, agrees).

It is contingent whether a given sentient should include the rule 'trap opponent's king' in its set of chess-playing rules. A rule-following computer could be built with hardware powerful enough to learn the formal rules of chess but scanty enough that almost no room was left over for rules of strategy; no doubt a sentient of like chess capacity is causally possible, so that the sentient could play chess but absorb so few rules of strategy that the rule 'trap opponent's king' would be crowded off the optimally successful list; e.g., perhaps there would be room for nothing closer than 'take opponent's pieces'. Likewise, perhaps

there is a low level of human intelligence at which the human could absorb so few moral rules that the rule 'maximize utility' would be crowded off the optimal list. However, it is prohibitively unlikely that a sentient of normal human intelligence could absorb so few chess-playing rules that the rule 'trap opponent's king' (or some logically equivalent expression) would fail to appear on the optimal list; how could such a human, not even knowing the object of the game, beat a human who did know? Likewise, it is prohibitively unlikely that a sentient of normal human intelligence could absorb so few rules for living in accordance with utilitarianism that the rule 'maximize utility' would fail to appear on the optimal list; how could such a human, not even knowing the object of the lifestyle, live it better than a human who did know?

Recall that the appearance of the rule 'maximize utility' on one's list does not preclude the simultaneous lower-level listing of 'refrain from maximizing utility' as a means to executing the former rule. Like the fish-spearer throwing the spear away from where the fish looks to be, the utilitarian may refrain from doing what they think would maximize utility because they doubt that doing that act would maximize utility. Just as the intelligent fish-spearer can use the rule 'spear away from where the fish looks to be' more efficiently with than without also having 'spear fish' on (a higher level of) their list of rules, the intelligent utilitarian needs the rule 'maximize utility' if they are efficiently on occasion to refrain from doing what they

think would maximize utility.

Regarding the question whether and how to use the rule 'maximize utility' to further the moral ideal of maximizing utility, CU supplies all its human adherents with the ontic criterion rule, 'maximize utility', for use as an epistemic top rule, for two reasons.

The first reason is that as an empirical fact humans are in the long run happiest in a fairly egalitarian society with a free flow of information. In such a society it is difficult and expensive for CU rules or anything else known to more than a few people to be kept secret. How could one prevent the people of a society largely accepting and practicing the epistemic RU short moral code from finding out that 'maximize utility' is the ontic criterion of their moral code? Surely this moral code will promote education, including university ethics courses, in which this information would normally be taught. And if people in this society know their ontic criterion of right action, will it not occur to them to use it from time to time as an epistemic criterion, for efficiency as the top rule? Surely it will, for presumably this moral code will encourage its citizens to think creatively and look for new, more efficient ways of getting things done. Better, therefore, to accept that the utilitarian ontic criterion will be widely known and often used epistemically, and to plan for this inevitability by promoting customs and techniques for properly using 'maximize utility' as the top rule (like dealing with teen lust by promoting condoms

and masturbation rather than trying to suppress sexual knowledge).

The second reason CU supplies all its human adherents with the ontic criterion rule, 'maximize utility', for use as an epistemic top rule, is that use of this rule indispensably benefits human identification of acts that maximize utility, a particular instance of the benefit of free information flow. Society under CU retains some division of labor, so that some rules are relevant only to some occupations or social positions, so that it is unnecessary for any one person to learn all the rules, but anyone may know any rule they wish to.

To illustrate the benefits of efficient identification of right action, consider Brandt's suggestion of the possible top rule (remainder-rule):

'do act x iff a person who knew the relevant facts and had them vividly in mind, had been carefully taught the other rules of this code, and was uninfluenced by interests beyond those arising from learning the code, would feel obligated to perform that action'
(paraphrasing 1963, p. 133)

While it is an empirical matter whether utilitarians using Brandt's suggested top rule would do better than those using 'maximize utility' as their highest-level epistemic criterion, human history reveals so many horrendous events being started by people extrapolating from their moral intuitions in a sincere but

vague way uninformed by metaethical theory that it is more likely that utilitarians using Brandt's proposed top rule would do worse, as game theory predicts via the chess example. The road to hell is paved with intentions coherent with good ones but vague and untested themselves, and invented by sentients evolved to have a significant capacity for both selfishness (Hardin 1985, p. 132) and self-deception (Trivers 1985, ch. 16).

Consider a sample advantage of the epistemic criterion of utility maximization over other possible top rules. Since this criterion is also the relevant ontic criterion, it is uniquely useful for quickly identifying objects with high probabilities of helping constitute states of affairs falling under the jurisdiction of moral rules. A human chess player aware that trapping opponent's king is the object of chess will be well able to spot the advantage of monitoring the defensive capacity of opponent's king and surrounding pieces, whereas an otherwise similar player unaware of the importance of trapping opponent's king would be significantly more likely to spend resources monitoring situations elsewhere. Similarly, a human moral agent aware that maximizing utility is the object of moral activity will be well able to spot the advantage of monitoring the welfare of sentients affected by its actions, whereas an otherwise similar agent unaware of the importance of maximizing utility would be significantly more likely to spend its resources monitoring other events, e.g., soap operas or sports programs.

Another advantage for humans of the epistemic criterion of

utility maximization used as a top rule is its help in interpreting other epistemic criteria even in the absence of conflict. As Diggs says, the follower of instrumental rules is often expected to know the goal to which his rule-directed action supposedly contributes--to know "what he is doing" in this sense. Not always, to be sure, but often he could not make a sound judgment of when and how to apply the rule without this knowledge. (1978, p. 214)

Though unrecognized by Mill (1861, p. 33), this advantage is significant; imagine trying to teach someone a canoe paddle-stroke without telling them that the point of the operation was to push water rearwards. Brandt agrees that his top rule should perform this function (1963, p. 133), but his suggested rule seems inadequate: by analogy it would advise the student of paddle-strokes to:

'move the paddle thus iff a person who knew the relevant facts and had them vividly in mind, had been carefully taught the other strokes of this paddle, and was uninfluenced by interests beyond those arising from learning the strokes, would feel inclined to move the paddle thus'

A canoe-paddler taught in this way may well get across the lake, but is unlikely to set a speed or safety record.

Thus while Brandt does implicitly accept 'maximize utility' as specifying utilitarianism's ontic criterion of right acts, he does not think it obvious that this ontic criterion is also the most efficient top rule epistemic criterion for humans. CU holds that this efficiency is safely inferrable from currently available data (Hare agrees - 1981, p. 50). Thus the difference between CU and Brandtian RU is minor.

CU does not differ greatly from (J.J.C.) Smartian AU either: it differs chiefly in holding that some rules should be followed almost exceptionlessly rather than merely used as guidelines, even when there's time to think carefully about the consequences of one's actions. Rules such as 'keep your promises', 'provide for your children', and 'don't punish the innocent' should be followed more regularly than mere rules of thumb or rough guidelines.

One game-theoretic justification for following such rules almost exceptionlessly even when there's time to think carefully is that some such rules are like 'drive on the right' in that one of their chief benefits is their tendency to coordinate cooperative behavior. The exact nature of the rule is less important than that everybody follow the rule. Some similar rival version of the rule might have worked just as well, *mutatis mutandis*, e.g., 'drive on the left' rather than 'drive on the right'; likewise, rules such as 'provide for your left-hand neighbour's children' are just as intrinsically morally acceptable as 'provide for your children'; as long as all

children are properly cared for by some adult, and the work is shared around, everybody's happy and morality is satisfied. For various contingent reasons of psychology, the latter rule is more likely to endure in a society than the former, hence is extrinsically morally better. The past enduring presence in a society of one rule from a set of such intrinsically equally efficient rivals is evidence that the rule will continue to endure. The rational moralist will therefore endorse and follow whichever rule of that set of rivals is predominant in local society, not worrying too much about which rule happens to predominate. Since the rule works best when followed exceptionlessly, and since there is no need to worry about alternative rules, the utilitarian follows the rule every time, even when there is plenty of time to think about alternative action, save in rare emergencies.

Another game-theoretic justification for following such rules almost exceptionlessly even when there's time to think carefully is that some such rules, e.g., 'keep your promises', are like 'use tit-for-tat strategy in iterated prisoners' dilemmas' in that they work best when followed exceptionlessly, and can be proven to be more efficient mechanisms of desire satisfaction than their rivals, so that again there is no point in thinking the matter over afresh each time one encounters the relevant situation. The institution of promise-keeping among normal humans always increases happiness over the lack of such an institution. The institution works best when promises are always

kept save in rare emergencies, mainly because while some lapses would be tolerable, we have little idea where lies the threshold between tolerable and intolerable lapses, and little idea how our lapses would affect the tendency of others to lapse (see Lyons 1967, p. 74). Once one gets in the habit of promise-breaking, one (and one's partners in promising) regularly lose benefits regularly secured by other sets of promise-keeping partners. The rational moralist will therefore know that there is nothing to gain by rethinking the matter every time a promise comes due. Thought is only necessary when deciding whether to make a new promise to someone with a spotty record of promise-keeping to you. Once a promise is made, it is always right to keep it save in rare emergencies. Such solemn promises must be distinguished from joking promises, only lightly-binding promise-like assurances, and so forth; this is most efficiently done by means of ritual oaths which by social convention bind when uttered, e.g. 'I swear on my honor to do x'.

The option of breaking a lower-level epistemic moral rule (such as 'keep your promises') in rare emergencies must be included partly to handle the catastrophic-alternative objection to demanding that any such rule be followed exceptionlessly, and partly to allow agents to secure the occasional large windfall benefits that help make possible the pleasant life that sentients should have, but which sometimes can be had, due to windfalls' arising unpredictably, only at the cost of breaking a lower-level rule.

Regarding the first of these two considerations, the catastrophic-alternative objection, it must be conceded that sometimes lower-level rules must be broken. Any plausible utilitarianism must make this concession to the complexity of the world, for history shows that sometimes catastrophe is the only alternative to breaking the lower-level rules of commonsense morality. For instance, it is part of common-sense morality, a part that presumably would find rational expression as a lower-level utilitarian rule, that cannibalism is wrong, that humans should not use human bodies as food. Yet survivors of air crashes in remote areas have sometimes found it necessary to eat human bodies in order to avoid starvation. Community moral intuitions tend to endorse such action. When rescued, such survivors generally receive sympathy rather than censure. As Diggs says of instrumental rules, "it not only makes sense to speak of its being proper to violate a rule, "successful violations" tend to be commended" (1978, p. 215). Such violations make sense as applications of the top rule, 'maximize utility', to adjudicate on difficult cases falling under lower-level rules.

Thus those who hold that "fairness demands that there be no exceptions" (Lyons 1967, p. 163) to moral rules are simply woolly-minded or fanatical, for if one takes their words literally, then one must infer that they would march innocents off cliffs for the sake of exceptionlessness. If fairness does require such fanaticism, then it cannot be the only or highest

value. A utilitarian version of fairness, interpreted as an epistemic criterion of right action that is widely useful in practical life, is more plausible, for it admits of exceptions.

However, to prevent abuse of this rare-emergency loophole, human moral agents need training and guidance in the use of their code's ontic criterion as a top rule:

Although it is difficult to specify conditions in which the violation of an instrumental rule is proper, surely the bare fact, "that by doing so one can better promote the goal," is not sufficient. The rule follower is not the sole or final authority on the propriety of breaking a rule, even when it is for the benefit of the other party. (Diggs 1978, p. 214)

Many possible sources of human abuse of 'maximize utility' as the utilitarian top rule could be acceptably minimized by the development of social norms for such use. Emergencies may be rare, but their general natures are foreseeable, hence may be discussed and planned for. To extend Mill's nautical analogy (1861, p. 31), sailors encounter big storms only rarely, and have at such times to act quickly in ways that greatly affect whether their fellow mariners live or die, sometimes having to choose between acting in accord with a high precept of ordinary good sailing, and acting to save many other people. Even so, sailors as a group can build up a tradition of knowledge to deal with

such crises, and as individuals become experienced at using this knowledge during crises. There is no reason to think that the general method sailors use to do this could not be equally well applied to dealing with rare moral crises (see Hare 1981, p. 52).

Regarding the second consideration, that agents must be free to secure occasional large windfalls at the cost of breaking a lower-level rule, again the utilitarian calculus and commonsense morality unite in endorsing such action. The calculus makes it plain that a sufficiently large windfall will greatly outweigh the modest harm a single violation of a lower-level rule tends to cause. Potential windfalls are not circumstances in which catastrophe looms, in the sense that a large amount of pain could be avoided but will occur as things stand; rather, they are circumstances in which a large amount of pleasure could be realized, but will not as things stand. In a catastrophe-avoidance emergency, prompt action is needed to avert pain; in a windfall-acquisition emergency, prompt action is needed to gain pleasure. A utilitarian moral code should try to handle cases of catastrophe-avoidance and windfall-acquisition roughly similarly, for according to the utilitarian calculus pleasure and pain are equally components of utility. The equal importance of pleasure and pain is reflected in utilitarian supererogation (discussed in detail in ch. 5), in which agents are praiseworthy for doing good deeds beyond what is morally required; the good deeds may be either the creation of pleasure or the prevention of pain.

Windfall-acquisition emergencies are like catastrophe-avoidance emergencies in that a little rule-breaking gains a lot of utility, so since it is, as Diggs suggests, justified to break a moral rule in a catastrophe-avoidance emergency (e.g., to break the moral rule of obeying the law of the land, by safely breaking the speed limit to drive a stroke victim to hospital), it would also seem justified to break a moral rule to secure a windfall. It might be objected that: on the one hand, we condone rule-breaking in catastrophe-avoidance emergencies because while rule-breaking tends to cause a little long-term pain, a lot of pain will occur anyway if prompt action is not taken; thus pain is unavoidable though minimizable. On the other hand, the rule-breaking to secure a costly windfall will cause pain that was avoidable; if the rule were followed, the windfall of pleasure would be lost, but no pain would occur. Since, for contingent reasons of human psychology, it is extrinsically more important not to cause pain than to create pleasure (Sikora, unpublished), there is a utilitarian reason to forego costly windfalls that does not apply to emergencies. Similarly, supererogation does not involve rule-breaking, so does not involve the pain that rule-breaking tends to cause, so does not run afoul of the utilitarian reason to forego costly windfalls.

Nonetheless, inspection of everyday human attitudes and behavior suggests that if the windfall is big enough and the rule-breaking cost small enough, we endorse the rule-breaking

behavior, and may even rebuke an agent who refrains from breaking the rule. For instance, if an unemployed person with children to support passes up an unexpected offer of a great job, in order to keep a promise to meet a friend for coffee, then even the friend is likely to chide the promise-keeper. Thus the utilitarian reason mentioned, that counts against costly windfalls but not against catastrophe-avoiding emergency rule-breaking or supererogation, seems to have the moral effect not of making the pursuit of costly windfalls irrational given utilitarian premises, but merely of setting the amount of utility gained (in the form of windfall pleasure), necessary to justify the costly windfall rule-breaking, as larger than the corresponding amount of utility gained (in the form of pain avoided), for catastrophe-avoidance emergency rule-breaking. This difference balances the subtle long-term harm that acquiring costly windfalls causes by creating avoidable pain.

These two considerations in defense of emergency rule-breaking represent the two extremes on the continuum of all possible utility packages to be avoided or acquired in emergencies. Catastrophe avoidances are avoidances of much pain, and windfall acquisitions are acquisitions of much pleasure. In practice, the utility packages to be acquired or avoided in emergencies are seldom composed of either pure pleasure or pure pain. Usually they are composed of a mixture of pleasure and pain, though so preponderantly of one or the other as to precipitate an emergency. In the case of a mixed catastrophe

avoidance, the utilitarian calculus unproblematically deems it worthwhile to avoid a little pleasure in order to avoid a lot of pain. In the case of a mixed windfall acquisition, the utilitarian calculus adds the small quantity of pain in the utility package gained to the small quantity of pain caused by the rule-breaking, and again deems the sacrifice worthwhile if the ratio of pleasure gained to pain endured is sufficiently great.

It might seem that there is a further non-emergency class of cases in which the breaking of moral rules is permitted, namely those cases in which only small amounts of utility stand to be gained, but the agent is justifiably sure that this utility would outweigh the utility lost by the rule-breaking. However, there are two objections to this view. First, as discussed in more detail elsewhere in this chapter, the small amounts of utility thus gained would be outweighed by the utility lost by other, less able agents being unjustifiably sure of gaining utility in similar situations in which utility would in fact be lost. Second, there is no need for such a practice, for there already exists a type of weak, flexible moral rule that covers such situations adequately.

This type of rule is the moral equivalent of the legal rule against jaywalking. Such rules are available to govern situations that are usually unimportant but sometimes very important, even potentially lethal. The important subsets of these situations are impossible to pick out accurately by

expressions concise enough for ready human use, hence must be covered by simple but greatly overreaching expressions. It is for contingent psychological reasons easier for humans to assimilate the more detailed criteria of the rule's use not as linguistic expressions but as habits of application. Thus we remember the linguistic rule 'don't jaywalk' but break this rule very often, e.g., when traffic is light and no crosswalk is handy.

Jaywalking-type rules function similarly to proverbs such as 'many hands make light work'. Proverbs are easy to remember and apply, but if taken literally and applied exceptionlessly would be hopelessly awkward and disruptive, as evidenced by the contrary proverb 'too many cooks spoil the broth'. These two sample proverbs coexist despite formal inconsistency precisely because each can be safely ignored in most of the situations to which it literally applies. The detailed behavioral criteria of use for jaywalking-type rules can be verbalized if necessary, e.g., when one is explaining one's actions to a traffic cop, but normally are not verbalized. A moral rule of this type is 'don't speak ill of the dead'. Unlike the stronger moral rules discussed earlier, jaywalking-type moral rules should not be followed with almost no hesitation - rather, a good deal of hesitation is desirable.

Given that there are various kinds of rule that a utilitarian ought to follow with almost no hesitation, and that humans are weak-willed and suggestible at times, it follows that

utilitarians should encourage humans in their society to regard such rules with respect, even to venerate them somewhat, and ritualize their performance. Such cultural practices are effective in encouraging willing human compliance with such rules. Venerating such rules makes breaking them in rare emergencies stressful, but this disutility is overall slight precisely because emergencies are rare; since veneration makes it easy and natural to follow these rules on other occasions, on the vast majority of which utility is gained, veneration short of fanaticism is helpful. Rules treated in this respectful way are therefore being used as more than mere rules of thumb or rough guidelines. Some rules therefore play stronger roles within utilitarianism than Smartian AU suggests is rational. However, since Smartian AU is perfectly consistent with rule-use per se, and since CU's improvement in the details of utilitarian rule-use follows from the application of game-theory to the task of utility-maximizing in human society, an application J.J.C. Smart endorses (1973, p. 57), the difference between CU and Smartian AU is minor.

(More precisely, I term this a minor difference because it is a difference at a low theoretical level, merely on the details of application of metaphysically well-understood higher principles to which both sides agree; hence, the disagreement would seem to be easily resolved. Only a little empirical spade-work is required, much as in a disagreement between two chess players about whether one should castle early in the game

as a matter of course, almost exceptionlessly, or whether one should castle early in the game only after thinking how this fits into the rest of one's game plan as constrained by the development of one's position thus far. The answer in chess turns out to be similar to the answer to the moral question. Weak players should castle religiously, and strong players should play it by ear. Similarly, agents of merely human intelligence should follow the relevant moral rules religiously, and agents of sufficiently superhuman intelligence should use the relevant rules as rules of thumb).

Thus Lyons' claim that rules endorsed by AU are either trivial or not rigorously generated (1967, p. 119-20), interpreted as an objection to CU (as Lyons intends - p. 145), trades upon the fact that the generation of CU rules, which uses ontic AU as a premise, also requires as premises some contingent facts about the psychology of the agents whom the rules will be used by and about the habitat these agents live in. Without these facts only trivial rules follow (e.g., 'lying is wrong except when it gains utility'), but with these facts some important high-level rules follow, as shown above. The identity of the top rule is shown to be 'maximize utility', and rules such as 'keep your promises' are shown to be worth following without exception save in rare emergencies for which identifying criteria can be taught. Such moral rules are as well-evidenced as chess rules such as 'trap opponent's king', 'castle early in the game', 'advance weak pieces before strong pieces', and 'control the

center of the board', hence these moral rules are generated as rigorously as any useful noncontingent rules can be. It might be held that nonetheless rules of this general type do not deserve the label 'rigorous', but since such rules, in chess and in ethics, achieve their objectives when followed and have rational game-theoretic justifications, this linguistic deprivation lacks sting.

Lyons also objects that ethics needs non-utilitarian conditions for justifiably breaking rules such as 'keep your promises', for if only the utilitarian "condition had to be satisfied, we would be justified in breaking rules much more often than we normally think we are" (1967, p. 186). However, Lyons offers no empirical evidence for this contingent claim, and given the frequency with which people in fact normally do break promises, presumably thinking in most cases that they are justified in doing so, this accusation is harsh. I can offer no direct empirical evidence to the contrary, but am strongly of the opinion that indirect evidence, by arguments too lengthy to go into here, suggests that the reverse is the case, so that if a CU moral code were generally accepted by humans, fewer promises would be broken.

Another objection to the justification of epistemic RU by ontic AU is that while ontic AU would admittedly entail that most humans should use RU for their epistemic moral code, it (allegedly) would also entail that exceptionally intelligent humans should not be encumbered by epistemic RU, for their

superior abilities at spotting beneficial ways of diverging from the conduct required by epistemic RU would thus be wasted. Such geniuses would therefore be justified in adopting personal moral codes, secret as far as possible, diverging from epistemic RU.

However, the problem of self-serving rationalization by individual code-choosers would remain for this view, for intelligence is a characteristic possessed independently of basic human emotions. Geniuses are as likely to be tempted to act selfishly as most other people. Moreover, some nongeniuses think that they are geniuses, and would choose individual moral codes without being able to thus behave significantly closer to perfect utility-maximization than they would have under ordinary epistemic RU. Meanwhile, the gains to be had from the use of exceptional moralities by real geniuses would be much smaller than might first appear. For instance, since these geniuses would still have to support epistemic RU for general public use, and since public subscription to epistemic RU would be undermined by public observation that geniuses were in practice often ignoring epistemic RU, the moral code geniuses publicly advocated, geniuses would therefore have to publicly obey epistemic RU.

If geniuses were readily distinguishable from other humans, then some aspects of this problem would disappear, for geniuses would not have to keep their individual moral codes secret. For instance, some loss of utility due to unpredictability would be avoided, for no one would expect geniuses to act just like other

people - they could be predicted to behave unpredictably, without nongeniuses' behavior also being unpredictable. However, geniuses are in fact indistinguishable from other people in all surface characteristics, so this would not work; hence geniuses would have to keep their individual codes secret.

Another reason why the opportunities for diverging from epistemic RU would be rare, and the resulting gains small, is that CU, by its explicit inclusion of 'maximize utility' as the top rule, allows a highly intelligent human to act properly in any emergency that this person's unusual gifts allow them to detect, without transgressing the moral code. Because all people in the society are trained to deal within the moral code with emergencies, rationalization respecting emergencies can be kept to an acceptable minimum.

Thus opportunities for geniuses' esoteric moralities to diverge from epistemic RU would not be during emergencies, when the potential gains would be great, and in any case would be rare, for the reasons given earlier. Rare, individually small gains add up to not much - a quantity easily outweighed by the disadvantages mentioned.

Since a form of utilitarianism, CU, the form assumed by OCU, can thus be defended that combines the main philosophical virtues of AU and RU while retaining the vices of neither, the third objection to OCU is refuted.

Chapter Five

SUPEREROGATION

1. Small beneficial sacrifices

The fourth objection to OCU is the objection from supererogation.

A supererogatory act is, to define it in as much detail as is here needed, an act, beyond the call of duty, that benefits someone other than the agent, and is altruistically motivated. Saints and martyrs are extreme examples of the kinds of people to whom supererogatory acts have traditionally been attributed, though at the other end of the spectrum, very small altruisms may be supererogatory, e.g., taking a stray cat to the SPCA. In the context of OCU, a supererogatory act is an altruistically motivated act, uncalled for by any rule in the short moral code of the agent's society (i.e., beyond the call of duty, hence, in a sense defined below, not obligatory), that both increases overall utility and increases utility for other sentients. The agent typically sacrifices some of its own utility to make these increases, but self-sacrifice is not essential to supererogation.

The fourth objection to OCU is that it has the unworkable feature of deeming some acts supererogatory. OCU's form of

supererogation can be criticised in three ways, some ways applying to all forms of supererogation, other ways applying only to some forms, including OCU's.

The first criticism, which might be termed the objection of small beneficial sacrifices, is that under this definition of 'supererogation', any personal sacrifice, no matter how small, for the sake of any gain for others, no matter how great, would count as supererogatory as long as it increased utility but wasn't called for by the short moral code of the agent's society; by the intuitions of some, that doesn't seem right (Sikora, personal communication). However, my proposed definition of 'supererogation' agrees with others in the literature in this regard, so is intuitively acceptable at least to some. My definition agrees with other definitions in the sense that each of these other definitions conjoined with OCU entails that any personal sacrifice, no matter how small, for the sake of any gain for others, no matter how great, would count as supererogatory as long as it increased utility but wasn't called for by the short moral code of the agent's society. For instance, Montague writes that

x is supererogatory for y (the person performing x) if and only if x is not required and y is praiseworthy for performing x. And given the nature of praiseworthiness, y is praiseworthy for performing x only if x has moral value and is not prohibited. Some

little benefactions are neither required nor prohibited, but are praiseworthy if performed with appropriate motives and beliefs. Hence some little benefactions . . . are supererogatory for their agents. (1989, p. 108)

Making a very small, uncalled-for-by-the-rules sacrifice for a very great gain for others is analytically not required by the rules. Since such sacrificial action gains utility, it has moral value, hence, in at least some cases let us suppose, is not prohibited by the rules of the short moral code. Since such action has moral value and is not prohibited, it is praiseworthy by Montague's definition. Since such action is praiseworthy but not required, it is also by Montague's definition supererogatory.

Thus I use 'supererogatory' to pick out a region of a continuum; anything over the relevant boundary, no matter how slightly over, is strictly speaking supererogatory. Analogously, near-tepid bathwater may still be on the hot side of the zero-point of the intensity continuum exhausted by hotness and coldness. The boundary OCU uses is the conjunction of the zero-point of the quantitative continuum of overall utility gained, and the zero-point of the quantitative continuum of utility gained for other sentients. If both boundary conjuncts are crossed, no matter how slightly, so that both quantities are positive, no matter how slightly, then the altruistic act in question is supererogatory. The amount of utility accruing to

the agent is irrelevant. It is typically negative, but may also be zero or positive. For instance, the agent may enjoy helping others, enough that the utility gained by the agent equals or outweighs that lost in the doing of the deed, e.g., the mailing of a small check to a charity. (At least, these quantities of utility are the ones relevant to our initial thinking on supererogation, when we take a God's-eye view of events; later, qualifications will be added to take account of the epistemic problem that acts do not always yield the quantity of utility that seemed probable by a given agent's evidence).

Another definition of 'supererogation' agreeing with mine in the relevant respect is Langford's definition:

works of supererogation . . . comprise the kinds of actions that it is generally held that we ought to perform for our close friends, but that it is not generally held that we are normally obliged to perform for our neighbour or a stranger, but which are done, nevertheless, for a neighbour or stranger. (1988, p. 442)

It is surely required by OCU's short moral code that, if the opportunity arises, one make a very small sacrifice to secure a very great benefit for a close friend, yet not required that one do this for a mere neighbour or stranger. For example, if an elderly close friend is moving, then (if I am healthy) I am remiss if I do not offer to help shift the larger items, though

not if I notice that a similar mere neighbour or stranger is moving, and make no offer. While it is doubtful that every act deemed supererogatory by OCU is of a kind that one ought to perform for one's close friends, it is very plausible that all acts involving very small sacrifices for very large gains are obligatory for close friends (since one has few close friends, so will not be overburdened), yet are merely permissible for strangers and suchlike (whose numerousness would overburden the agent often enough to prevent the relevant rule from enduring very long in the moral code).

Similarly, Dancy states that:

supererogatory acts are those that, though they have merit or value, still lie beyond the call of duty. At the limit, these are the actions of saints and heroes, actions that, though of supreme merit, cannot be said to have been the agent's duty. There are also supererogatory actions of less exceptional value; quite ordinary actions can exceed the demands of duty, while still attracting approval. (1988, p. 173)

I shall not argue the point in detail, but this definition also seems to agree with OCU's on the salient point.

Thus there is among those who have studied the subject significant intuitive support for the view that any personal sacrifice, no matter how small, for the sake of any gain for

others, no matter how great, would count as supererogatory as long as it was gainful on balance but wasn't called for by the relevant moral code. Applied to OCU, this would pick out suitable acts that weren't called for by the short moral code for the agent's society. It is therefore incumbent on those who find this view of supererogation unintuitive to develop a more explicit theoretical objection to it. Meanwhile OCU assumes the view, and incorporates it into the definition of 'supererogatory'. Thus the objection of small beneficial sacrifices fails.

2. Valuable but not obligatory

Some hold that the practice of supererogation is unjustified because the concept of supererogation is paradoxical, in that a supererogatory act would make things better all things considered, but is not obligatory. This nonobligatoriness seems inconsistent with

the resilient intuition that if an action is good, this gives us a reason to do it, and if it is the best available, we have more reason to do it than to do anything else and so ought to do it. If this intuition is sound, the best available action will always be the right one, the one that one ought to do.

(Dancy 1988, p. 176)

This general criticism of any morality's form of supererogation is starkly applicable to OCU's, for according to OCU all and only utilities intrinsically ground reasons for action. Call this criticism of supererogation the 'justification objection'. One might save OCU from this serious metaethical objection by abandoning supererogation, but this seems too high a price. Supererogation is familiar enough as a feature of benign commonsense moralities around the world that we may speculate that it is in some way justified. I shall argue that certain evidence renders this speculation plausible. Kagan writes:

according to Heyd, the justification of supererogation has two "aspects": the negative aspect shows the justification for limiting what is morally required; the positive aspect shows "the value of non-obligatory well-doing as such" (1984, p. 241)

I shall follow this program of justification.

The negative aspect of OCU's justification of supererogation is the aspect of limiting what is morally required by showing that the short moral code optimal for human societies makes obligatory only a proper subset of the many valuable acts open to a typical human agent.

This objection to OCU's form of supererogation, that it cannot so limit obligation, might be termed the objection of

pervasively many obligations (adapting Montague's phrase - 1989, p. 104), and is due to Baier (1958, p. 203-04). It is the objection that since it is obligatory for the utilitarian agent to act on every occasion on which it has the opportunity to increase the world's utility, the utilitarian agent will find its entire waking life dense with obligatory acts. Duty would pervade the morally obedient agent's life to the exclusion of frivolity and relaxation. In McCarty's words,

the fantastically virtuous person rises daily
with moral determination, eats with moral
discrimination, organizes all activities
toward moral destinations, and retires in
moral deliberation. (1989, p. 44)

OCU's response is essentially that rules making so many acts obligatory would seldom be obeyed, hence that such rules would tend to come into disrepute, and gradually get dropped from the moral code, both because of fading by neglect, and because of active defensive suppression by moralists since flouted rules tend to bring morality in general into disrepute. The details, however, are complicated.

For instance, the rules in question are divisible into two classes, flouted for different reasons. J.J.C. Smart's response to the objection of pervasively many obligations is effective in the large first class of cases:

Baier holds that (act-) utilitarianism must
be rejected because it entails that we should

never relax, that we should use up every available minute in good works, and we do not ordinarily think that this is so. The utilitarian has two effective replies. The first is that perhaps what we ordinarily think is false. Perhaps a rational investigation would lead us to the conclusion that we should relax much less than we do. The second reply is that act-utilitarianism premisses do not entail that we should never relax. Maybe relaxing and doing few good works today increases threefold our capacity to do good works tomorrow. So relaxation and play can be defended even if we ignore, as we should not, their intrinsic pleasures.

(1973, p. 55)

J.J.C. Smart's reply is sufficient to defeat the objection of pervasively many obligations in all cases in which doing good works for others would cost the agent significantly, e.g., in time and effort. While many people have more utilitarian obligations than they fulfill, they do not have so many as to mire them in ceaseless work.

In the second class of cases, where little or no sacrifice of utility, e.g., via time and energy, would be needed from the agent for each of the very numerous acts in question, OCU holds that, nonetheless, agents would not be obliged by the short moral

code to do all these acts. For instance, the billionaire need not be obliged to sign a check for charity, even though the act would take only a few seconds and would help sentient considerably. A rich person would no doubt be required by the short moral code to give some amount to charity, but beyond that amount, even though the person had a great deal of money left, donation would be supererogatory. It would be supererogatory because a rule making such action obligatory would not be obeyed enough to endure, hence would be deleted from the moral code. We could logically possibly do that many things, but would not actually do them, for the causal fact is that humans are seldom so generous as to disperse their worldly goods like snowflakes. For instance, as the fall of communism showed, people deprived of the fruits of their labor cease to labor, even when the seized goods are dispersed to needy people. Therefore, as with the first class of cases, the obligations thus ignored would be brought into disrepute and disappear from the moral code.

It might be counterobjected to OCU's reply to the objection of pervasively many obligations that in the long run it would not bring moral rules of obligation into disrepute to make the relevant acts morally compulsory rather than morally optional, for in fact humans, with proper preparation and training, could and would act very much more diligently than they now do to promote utility. Consider, therefore, some more detailed points of psychology, which suggest that a moral code successfully committing its followers to relentless pursuit of duty would not

be a moral code conducive to utility maximization.

First, relentless hard work is causally incompatible with the balanced, thoughtful, reflective habit of mind, including an appreciation of pleasure and emotion, that leads one to be a utilitarian in the first place. John Stuart Mill and his well-known nervous breakdown form a typical example of a human mind in which both types of activity were encouraged by misguided educators (see Mazlish 1975, especially ch. 10).

The belief that happiness is the only goal worth promoting for its own sake is the central belief of hedonistic utilitarianism, so cannot be abandoned without abandoning the main premise intellectually supporting the view that one should work ceaselessly for others' happiness. Yet without proper intellectual support, the belief that one should work ceaselessly for others' happiness cannot itself survive - it gradually changes into a related belief able to derive intellectual support from non-hedonistic beliefs. Anthropology reveals that cultures in which people work very hard are not hedonistic but disciplinarian, not individualistic but conformist. Happiness does not thrive in such a society.

People tend to seek happiness for themselves; just as photosynthesizing plants are phototropic, humans are happiness-tropic. To get people by working ceaselessly to act against their drive for happiness, a pro-work belief must be maintained. This can only be done by threat of some worse alternative, e.g., hunger or torture or death, or by persuasion

that work is better than self-happiness, e.g., by means of cultish organizations glorifying work. Some examples that come to mind are puritanical Protestant religious sects, the Japanese sociopolitical unit, and the German traditional nationalist movement, with its slogan 'Arbeit macht frei' (= 'work makes one free'). The threat of a worse alternative produces a siege mentality; just as many drafted soldiers in continuous combat experience battle fatigue after a while and cannot function, previously normal people who work ceaselessly tend to crack up after a while. People can be trained to work very hard, but must be raised in that mentality from childhood. Such people must be raised to be psychologically stunted; too much education will cause them to ask questions awkward to the cultish organization doing the training, and too wide a variety of emotions will cause longings and dreamings incompatible with ceaseless work. Psychologically stunted people en masse form societies that are rigid and insensitive to the needs of individuals, particularly the weak or mute; such societies contain much needless suffering and needlessly forgo much pleasure. If the resources available to such a society suddenly increase, due to, e.g., trade access to a hitherto unexploited land, or the acquisition of scientific knowledge permitting better utilization of local materials, then the hard work such a society engages in may produce prosperity. However, belief in the value of ceaseless work does not thrive in a habitat of human prosperity; it is not a coincidence that as the three work-glorifying ideological organizations noted above

began to achieve prosperity, permitting education and spare time, their members began to lose their zeal for work.

Thus the objection of pervasively many obligations, respecting cases of supererogation requiring significant agent sacrifice, depends upon a false factual premise about the amount of altruistic good work human beings are on average capable of in the long run. Utilitarians agree, or ought to, that humans should do as much altruistic good work as they are capable of in the long run, but in fact most humans aren't capable of very much. Humans who work relentlessly hard tend not to do work that creates happiness, and humans who do work that creates happiness tend not to work relentlessly hard. There are sound scientific reasons for thinking that this situation cannot be changed.

Some humans are capable of both relentless hard work and ethical sensitivity, but they are very much in the minority. With appropriate psychological and genetic engineering, perhaps all humans could be made this way. However, it seems unlikely that a society of such humans would be stable. Relentless hard work may help us get to utopia, but once we are there it is unnecessary. Hi-tech humans, like humans in some hunter-gatherer societies, need not work more than 2 or 3 hours a day to keep society running. The bulk of the work performed by contemporary humans either is needless, irrational busywork, or is needed only to counteract the blunders of the irrational majority - but in utopia, irrationality would be rare among humans. The inhabitants of utopia would have no reason to maintain the

artificial controls required to sustain the capacity for relentless hard work, so being rational, would not maintain it (if they had ever begun it, which is doubtful).

Second, similar considerations of psychology apply to cases in which little or no time or energy need be expended to accomplish the supererogatory act. In a society in which many are poor, so that there is significant need for such acts, human competition for worldly goods will tend to be intense enough to keep human nature sufficiently selfish on average that moral rules making such generous action obligatory would not be obeyed very often. The moral rules in question would therefore fall into disrepute and disappear from the moral code. In a society in which few are poor, so that there is no significant need of such acts, there will hence be no significant need for moral rules making such generous action obligatory. Ordinary tax-supported programs will be sufficient to ensure that no one involuntarily goes without the basic goods of life.

These facts about human nature will be reflected in a rational utilitarian moral code by means of rules specifying the modest number of altruistic good deeds the average human should feel obliged to do. Individuals are permitted to do more - supererogation is not banned by utilitarianism - but individuals using this liberty should be warned of the possible harmful side-effects of supererogation, and trained in testing themselves for possible damage, much as athletes are trained in elementary sports medicine so that they can watch for possible harmful

side-effects of the unusual strains their bodies undergo. The harm to supererogators might be physical, such as when a chronic volunteer ruins their bodily health, or mental, as when a chronic donator of money loses their ability to say 'no', and sends checks to charity so often as to squander their savings for their old age, as many people have in response to televangelists' pleas on behalf of pseudocharities. No rule requires strong-willed humans to perform these supererogatory acts, even though they are capable of them without self-harm or harm to others by side-effect, because it is too difficult, hence costly, to identify these rare individuals among a population of ordinary humans. The extra cost involved would cancel the gain in utility otherwise resulting.

AU has received unwarranted criticism via the objection of pervasively many obligations, because it has not been widely appreciated that AU can make such moderating rules for citizens. RU is liable to the same technical objection, but has largely escaped criticism because it is easier to see that RU rules must be designed to endure and be generally obeyed in a society of ordinary, weak-willed humans. As we have seen, AU and RU are, properly construed, the same moral theory, CU, which hence is the beneficiary of all the intuitive support RU enjoys on this issue, as hence is OCU also.

The objection of pervasively many obligations thus dealt with in the abstract, we may now examine just what sorts of rule will bring about the appropriate rate of self-sacrificing action

in a utilitarian society. While it is clear that for the average human the sacrifices involved in much supererogation are too high for the gains made, in that the average human could not keep up this level of work unrewarded by personal happiness for long, and the gains made would be insufficient to warrant asking for the sacrifices anyway in order to get even a small measure of compliance, it might be asked just how the correct rate of self-sacrificing action is to be determined. The correct rate of self-sacrificing action depends upon the correct sacrifice/gain ratio, but what ratio is that? The straightforward answer is that the correct ratio is anything less than 1; that is, action in general is *prima facie* demanded by utilitarianism anytime a net gain in utility would result, and this applies no matter who gains and who loses. Thus if I would lose 100 net hedons (measures of happiness) by act A, and others would gain 101 net hedons, then I ought to do act A, in the sense that it would be better if I did act A, hence that the short moral code ideally would find a way of directing me to do act A. However, and the importance of this qualification cannot be overstressed, it is very difficult to tell whether this ratio obtains of a given act, and, because of obscure long-term considerations, many sorts of act of which this ratio appears to obtain do not in fact yield any net gain when performed by an average human. The straightforward answer given above is an ontic criterion, whereas for practical purposes we need to know what epistemic criteria identify the acts thus specified.

Since matters of general human behavior are, according to OCU, best governed by rules from a short moral code, we are thus asking what sorts of rule will, if generally followed, result in a sacrifice/gain ratio of less than 1. One obvious rule in the code will command citizens to do volunteer work for altruistic good causes at a specific regularity. For instance, humans might be commanded to do such work for a shift of 2 or 3 hours once a week. However, many of the rules involved - for it is the net effect of the whole code that matters - would be far from obviously connected to altruism. For instance, the general rule to obey the law of the land (save for conscientious objection to particularly bad laws) contributes enormously to altruism. In particular, payment of taxes according to tax laws enables the government to carry out all sorts of worthy social programs of most benefit to those who pay the least tax money (for their income is low); by the law of diminishing returns, net happiness is thereby increased. It takes constant reminding to maintain public awareness that despite government inefficiency, tax money is on average spent wisely enough to gain utility over leaving the money in the public's pockets.

Involvement in politics is also a duty that contributes greatly though indirectly to altruism. To illustrate: after the recent fall of communism in the Soviet bloc, an East German citizen visiting West Germany for the first time said something like, 'I thought that everyone here would be working very hard out of the profit motive - but they don't work any harder than

us, not hard at all. I don't understand where all the money is coming from.' Educated people know that the money comes from the efficiency of the whole system. The invisible hand of the market directs wealth to productive enterprises, the governmental social safety net prevents human talent from being wasted or destroyed by bad luck in personal circumstances, and far-sighted laws protect valuable objects of national heritage, such as forests, cultural treasures, and clean water and air, from being lost to the tragedy of the commons (at least, this is the ideal - no country has yet achieved perfection in these matters, though some come tolerably close in most respects). The value to altruistic causes of such a system is enormous, and the creation and maintenance of the system depend on the involvement of individual citizens in democratic government. Therefore altruism is greatly advanced by the rule that citizens should get involved in politics, and, so that their involvement is productive, the rule that citizens should educate themselves to the full degree of their intellectual talents.

The game-theoretic reason underlying these sorts of rule is that by dividing self-sacrificing altruistic behavior into small chunks and getting people to do them at regular intervals over a lifetime, the moral code can structure the giving and receiving of benefit by and to individuals as an iterated prisoner's dilemma, making cooperation rational. Individuals can observe each other's behavior and take action against those falling too far behind in their giving. Taken as a whole, an individual's

giving largely ceases to be self-sacrificial - it is mere enlightened self-interest (though not all citizens are enlightened enough to regard taxpaying and so forth in this way). One gives up a lot in, e.g., some promise-keepings or refrainings from theft, but over a lifetime an average human is more than compensated for such loss by the accumulated gains, caused by one's givings-up, of promises kept by others and by enjoyment of property that others do not steal.

There is a small class of humans for whom such sacrifice is not in their long-term self-interest; namely, humans selfish enough to be untroubled by abandoning other sentient to their own devices, and cunning enough to make their way in the world without obeying moral rules, including those moral rules underwriting legal laws. It is insufficient to respond that the system would collapse if everyone scoffed at morality and law, for not everyone will scoff at morality and law. OCU nonetheless requires such selfish yet cunning humans to obey moral rules anyway, for it is rational to help others whether or not the potential helper realizes that it is rational (as discussed in ch. 3). The blaming of such egregiates is in itself useless, for they tend not to change, but is an unavoidable part of the system of blaming that keeps most citizens well-behaved.

Thus appropriate, moderate rates of altruistic self-sacrifice, which nonetheless pick out, to the highest humanly achievable degree of accuracy (i.e., less than full accuracy), all and only acts for which the sacrifice/gain ratio

is less than 1, are brought about by utilitarian rules commanding humans to perform the sorts of action described above. Some such rules are transparently altruistic, some are not transparently altruistic. These rules ensure that the correct sacrifice/gain ratio is achieved because they are designed to maximize present happiness without stealing resources from the future (e.g., without using more than the present population's fair share of planetary resources, such as by exhausting a valuable substance that cannot be quickly regenerated, such as oil or the ozone layer, or by extinguishing a major species of organism, such as the blue whale or the passenger pigeon). Maximizing on a steady-state basis is favored over a pulse pattern of resource consumption by the law of diminishing returns. While there is the trickle-down effect whereby losses to resource-stealing by past generations can be outweighed by gains from resource-processing knowledge discovered by those generations, this effect works only in the early stages of a civilization, and is negligible in the succeeding billions of years. Thus rules for general use will run society in a steady state, ensuring that universal utility is maximized. As mentioned earlier, this is a kind of success that cannot be improved upon.

Thus the optimal moral code for a society would be short; the code's rules would not pick out more than a general framework of behavior for a society, for any framework more detailed would causally necessitate a code too complicated for the average human to absorb. Since it is not, in the relevant sense, obligatory to

do what is not commanded by the rules of the short moral code, it is not obligatory to do the acts OCU deems supererogatory.

I have briefly and informally introduced some relevant senses of 'obligation' (and, *mutatis mutandis*, interdefinable terms such as 'obligatory' and 'permissible'). To complete OCU's reply to the objection of pervasively many obligations, let me now formally distinguish the sense of 'obligation' in terms of which 'supererogation' is defined from two other senses useful to OCU and many related moral theories.

One's ontic obligation is to do an act that will maximize utility. One's epistemic obligation is to do an act that one's evidence suggests will maximize utility. One's rule obligation is to do an act of a type that is in conformity with the rules of the short moral code for one's society.

An ontic obligation is an ontic criterion of right action; it says what a right act is. An ontically supererogatory act, a beneficial altruism that goes beyond one's ontic obligation, is therefore impossible, as suggested by the intuition Dancy describes. One's ontic obligation is to maximize utility; no act could go beyond that amount of utility. The sample condensed derivations of some common moral terms given in §1.2 were derivations of these terms taken in their ontic senses:

for 'ought to do X' read 'would do the most good by doing X'; for 'is obligatory under any circumstances' read 'would uniquely do the most good'; for 'is obligatory all else

equal' read 'all else equal will do the most good' . . . ; for 'has a duty to do X' read 'all else equal would do the most good by doing X'.

Ontic obligations are indeed pervasive; action is partitioned exhaustively into sin and duty (though including, e.g., duties to relax and enjoy).

An epistemic obligation is an epistemic criterion of right action; it says how a right act may be known; that is, it is correlated with right acts, hence is evidence of their presence, yielding a reason for action. Any act picked out by such a criterion as ontically right is one's epistemic obligation. Some humanly knowable epistemic criteria are completely accurate; the acts they pick out are all ontic obligations (note that the agent need not be certain in order to have knowledge; the complete accuracy lies in the truth of knowledge, not the evidence of its justification). Humanly knowable completely accurate epistemic criteria are not rules, which if simple enough to be humanly knowable would encounter exceptions sooner or later in a complex world such as the actual world; instead, such criteria pick out acts by means of definite descriptions (see Russell 1912, p. 52) of particular acts, e.g., 'the casting of one's vote for candidate X in tomorrow's election'. Other humanly knowable epistemic criteria are not completely accurate; not all the acts they pick out are ontic obligations. Such criteria to be useful in acquiring moral knowledge must pick out a good weight of ontic

obligations, hence to be simple are usually rules, expressable via descriptions of wide application, e.g., 'a keeping of a promise one has made'. These criteria pick out enough ontic obligations of sufficient weight that the humans using the criteria do more good, even allowing for the harm their mistakes cause, than they would using other epistemic criteria, even more complex wholly accurate epistemic criteria, in their place in the short moral code (a claim closely related to Sikora's error theory discussed in §4.1). An epistemically supererogatory act, a beneficial altruism that goes beyond one's epistemic obligations, is therefore impossible. To be supererogatory, an act must be done intentionally; the agent must be aware that its act will gain utility and benefit another. But this awareness as knowledge requires evidence; since the agent's evidence in the present case suggests that the act in question is an ontic obligation yielding a reason for action, the agent possesses an epistemic criterion of the act's rightness. The agent may also possess other epistemic criteria, e.g., incompletely accurate rules, suggesting that the act is not an ontic obligation, but by hypothesis such criteria are in this case mistaken and epistemically outweighed by others of the agent's criteria. Thus any putative epistemic supererogatory act reduces to an epistemic obligation, hence does not go beyond one's epistemic obligations, hence is not really supererogatory.

However, rule supererogatory acts are possible. Rule obligations are all and only those acts commanded by the rules of

the moral code, which rules though based on utility may, as in the human case, be prevented by the relevant agents' average limitations of evidence or intelligence from commanding all and only ontic obligations or even epistemic obligations. It is thus possible for rule permissible acts to fail to maximize utility, and for the agent to possess non-rule epistemic criteria giving the agent knowledge of this failure by way of knowledge of other rule permissible acts gaining more utility. Rule supererogation is thus possible for agents in sufficiently subomniscient cultures, including all human cultures. Thus, in my discussion of utilitarian supererogation, by 'supererogatory' I mean, unless otherwise indicated, 'rule supererogatory'.

To my explication of OCU's version of supererogation it might be objected that I have merely equivocated on 'obligation', so that the objection is merely transferred to OCU's code of rules, which seems not to require some obligatory acts to be done. I reply that this construal of the objection depends on clinging dogmatically to the exact traditional intuitive sense of 'obligation', whereas the nature of reforming definism is to depart a little from traditional, intuitive usage (and to depart in order to more closely approach other intuitions relevant to ethics). The reforming clarificatory moves OCU recommends respecting the word 'obligation' and its attendant practices are part and parcel of the broad utilitarian program of discovering and implementing new socioethical practices that approximate traditional practices, allowing them to harness human emotions as

they are, yet that are pointed in utility-gaining directions (e.g., Hare 1981, ch. 2). Utilitarian interpretations of the social institutions, including reactive attitudes, of praise, blame, respect, contempt, guilt, shame, regret, remorse, repentance, punishment, rehabilitation, virtue, rights, justice, fairness, compunction, gratitude, forgiveness, esteem, and so forth, are arrived at in much the same way. One might, of course, merely enlarge the objection to apply to this whole program, but given the practical success of this program (for utilitarians flourish psychologically and are surely no worse than average behaviorally) combined with its conceptual coherence, it seems wiser to abandon the exact traditional intuitions in question. Some of the traditional senses of moral terms turn out to be incoherent when examined closely (e.g., Kantian 'freedom' and 'desert' - Sikora, unpublished). Moreover, some other prominent ethics besides OCU require much theoretical interpretation of traditional moral terms and practices - contractarian theory, for instance. If these ethics too must be rejected, then we are left with little of philosophical merit. If we must choose between blind traditionalism and philosophical analysis as metaethical approaches, surely it is more reasonable to choose the latter.

As it happens, OCU's defence of supererogation is closely related to its defence of praise and blame. Rather than treating praise and blame as items appropriate to agents for suitable action on the ground of Kantian desert, OCU follows the general

utilitarian solution of treating praise and blame as useful practices. Recall J.J.C. Smart's admonition:

I beg the reader . . . to bear in mind . . . Sidgwick's . . . most important distinction between the utility of an action and the utility of praise or blame of it. The neglect of this distinction is one of the commonest causes of fallacious refutations of act-utilitarianism. (J.J.C. Smart 1973, p. 56; see also Lyons 1967, p. 27).

Utilitarianism justifies praise by abandoning the traditional view that there is a deep Kantian metaethical link between agent and praise, and taking up the view that it gains utility to make a formal normative practice of encouraging morally successful agents on certain kinds of occasion by means of the language of praise; complementarily with blame. These are efficient ways of harnessing human emotions as they are and are likely to remain. Praise functions well as positive reinforcement of choice and behavior, yet for easy assimilation is couched in the language of human reactive attitudes (see Strawson 1982). There is some question how we are to make metaphysical sense of these attitudes given that all events, including human moral choices, are either determined or random. However, that is a problem for any moral theory, not just utilitarianism, and OCU, by locating all intrinsic value in hedonic items, renders itself independent of the existence of a deep Kantian basis for reactive attitudes.

OCU could accommodate such a basis if it existed, but since it seems not to, my discussion of OCU omits description of such an accommodation.

To be morally useful, praise should not necessarily be issued on every occasion on which an agent acts so as to gain or maximize utility. First, there are very many such occasions, so the institution of praising would be impractical if it required praise on every such occasion. Second, some but relatively few such occasions are positive (rule-) obligatory acts (as opposed to negative obligatory acts, such as not robbing, which are very numerous and usually trivial to perform). Since these rules specify the supporting framework of a morally optimal society, positive acts commanded by the rules are particularly worth positively reinforcing. Therefore it is plausible that praise be demanded or at least encouraged on such occasions, or on some significant percentage of such occasions (positive reinforcement seldom achieves maximum return on resources expended when doled out on 100% of successful trials, and even leaving the cost of reinforcement aside, performance is often improved by lesser percentages of reinforcement - Timberlake and Lucas 1989, p. 269 - by a sort of 'tease' effect). Praise is also useful now and then to reinforce negative action, but far more rarely - mostly on occasions of resistance to unusually strong temptation.

On the remaining occasions, on which utility is gained or maximized but the act is not obligatory, it is plausible that praise is permissible but not required, nor encouraged so

strongly as in similar cases where rules are obeyed by positive action. Praise is permitted partly to avoid cognitive dissonance in the potential praiser, since the act is known to gain utility, which is posited to be the only intrinsic value, and partly to promote the gaining of a little utility beyond what the rules specify, for a hedon more is a hedon more. Praise is not required or strongly encouraged, for fear of creating social expectations of beneficence beyond duty, which in turn would cause pointless guilt in the majority not doing such actions. Such guilt would be pointless because it could not usefully change their behavior; by hypothesis most humans could not live up to standards significantly higher than those of the short moral code.

Utility-promoting acts beyond what is called for by the rules are therefore praiseworthy, hence eligible for supererogatory performance.

To this account of supererogation it might be objected that it is actually harmful to praise acts beyond the agent's rule obligations, since the praise will encourage them to try to do what in the long run they cannot. However, agents typically acting supererogatorily by so acting demonstrate a supernormal quantity of moral perception and strength of will that makes it more likely than not that they could keep up good deeds at a pace above that prescribed in the short moral code for their society. As well, it is part of the convention that praise of a supererogatory act is praise of that particular act, not of the

practice comprehending the act. Agents typically declining supererogation by so doing demonstrate a suboptimality of moral perception and strength of will that makes it more likely than not that they could not have kept up good deeds at a pace above that prescribed in the short moral code for their society. These likelihoods rationally warrant both our not blaming those who decline supererogation, and our praising those who perform supererogation. Thus the acts picked out are both praiseworthy and not obligatory, in that the acts which are done tend to be praiseworthy, and the other acts which are not done tend not to be blameworthy; the potential acts though all falling under the same act-description, are separated by which are done and which are not. Therefore the objection fails.

It would take detailed empirical studies to prove that this type of praising function would optimally harness human emotions, but it is at least plausible on present knowledge - plausible enough to ground OCU's version of supererogation. While empirically unproven, this praising function has not been empirically refuted, is supported by some empirical evidence, and is metaphysically uncontroversial hence metaethically sound.

To Dancy's concern that any act better supported by reasons for action than its alternatives must be obligatory, OCU therefore replies that while this is true respecting ontic obligation and epistemic obligation, it is false respecting rule obligation. Rule obligations arise not only from reasons but also from the complex web of motivating institutions described.

A reason for an action, even a conclusive reason, is necessary but insufficient for the act's being a rule obligation. A reason for action may provide some motivation, but not always enough to get the potential act chosen for actualization over competing acts. The complex utilitarian system of rules, praise, blame, etc., is designed to shunt available motivation around so that in the long run the acts with the most compelling reasons tend to get chosen and done (thus acts are effectively chosen between even when they are not colocated alternatives, but distant in space-time). Acts with reasons to be done but little prospect of obtaining sufficient motivation to get done are not made rule obligatory, for there would be no advantage to doing so. The intuition to be opposed to the intuition Dancy mentions is that ought implies can. Without motivation one cannot will the act; hence without motivation it is not the case that one ought to will the act.

Thus the first half of the second criticism of OCU's supererogation, the justification objection, has been answered.

The second aspect of OCU's program of justification of supererogation is the positive aspect of showing the value of non-obligatory well-doing. In light of the explanations already given, this is easily shown: it has been admitted that the shortness of the code of rules prevents it from making all ontic or epistemic obligations into rule obligations, and surely some of these left-over obligatory acts are both ontic and epistemic obligations, hence utility-gaining and justifiably believed so by

the relevant agents. These agents then have the option of doing supererogatorily any of the many such acts that are permitted by the rules, and that gain more utility than other permissible acts. Since these acts gain utility, they are valuable.

Some acts that are utility-gaining and permissible but not obligatory by ordinary rules are nonetheless reasonable responses to emergencies, hence are rule-obligatory by the top rule 'maximize utility'. To this view it might be objected that if the top rule covers emergencies, then it covers supererogation too; just saying that it doesn't cover supererogation doesn't make it not cover supererogation. The covering or lack of it issues from the rule itself. The rule says to maximize utility, and the supererogatory act is known to produce more utility than its nonsupererogatory alternative. To this objection OCU replies that the human tendencies to rationalize, etc., discussed earlier, prevent the rule 'maximize utility' from justifiably being applied directly in cases of supererogation, though the rule does apply indirectly, e.g., in the background rationale for restricting its direct application. The agents' motivation for following this restriction is guided and amplified by praise and blame. This guidance is done by means of learned criteria and cointexts for rule-following. Humans have the causal capacity to learn criteria and contexts for rule-following, as shown by the difficulty humans have using new rules until they have learnt criteria and contexts, and by the unfortunate ease with which the process can be subverted by irrational criteria and contexts,

producing an immoral double standard. Human legal systems, with their close attention to often-obsolete precedents, offer many examples of both phenomena. Thus OCU's solution to the objection has been shown to be rationally justified, agent motivatable, and causally possible.

It might also be objected that in most cases where an agent by declining to act supererogatorily declines what seems like achievable utility, the utility was not in fact achievable; the short-term gain would be accompanied by outweighing long-term loss. The option of such agents doing such acts with no such associated downside is a mirage. Thus no supererogatory act was possible; there was no available act which would have gained utility, hence nothing valuable beyond the rules.

Analogously, for an alcoholic, the act of having one drink may seem more beneficial than the act of having no drink, but in fact is a mirage, for the act cannot be done independently. Having had one drink, the alcoholic would end up having many drinks, and the world would suffer a net loss of utility. The option of having just one drink was not open.

To this objection it can be replied that while in many cases no utility could have been gained, in fewer but still many cases significant utility could have been gained. Some achievable utility is lost by OCU. It sometimes happens that an able, strong-willed human knows that a utility-gaining action would benefit others, yet not benefit them enough to constitute an emergency. The act is permissible but is commanded by no rule of

the moral code, hence is supererogatory. On such occasions the agent would have sufficient reason for action, and would know it, yet might or might not find enough motivation to act ontically rightly. Similarly, a champion high-jumper, though well able to jump a certain height, may or may not do so on a given attempt. It is not the fault of the coach when the jumper fails; we may suppose that no other coaching method would improve the average performance of the jumper. Likewise, it is not a deficiency of OCU when achievable utility is lost on the occasions described above. It is, from the point of view of the moral code, simple bad luck. No further instruction issued by the moral code would improve the situation on balance. One can only hope that the agent will independently recognize the facts and act ontically rightly. On some such occasions the agent will have sufficient motivation; the agent will thus have a reason and motivation, but no obligation. On other occasions the agent will not have sufficient motivation; the utility thus lost is achievable in that though determined not to actually exist, it would have resulted from an act that the agent was free to perform.

Thus the second half of the second criticism of OCU's supererogation, the justification objection, has been answered; since both its halves have been answered, the objection as a whole fails.

3. The sacrifice/gain ratio

It might be objected that it is impossible that the resulting set of rules could ensure that the correct sacrifice/gain ratio is on average achieved by everybody, simply because there is no such thing as a sacrifice/gain ratio that is correct for everybody (Sikora, personal communication). This is the third criticism of OCU's form of supererogation; call it the sacrifice/gain ratio objection.

Non-cognitivist versions of CU are hard-pressed to answer this objection, for, as will be discussed in more detail in §6.1, non-cognitivist utilitarianisms cannot consistently claim (and tend not to claim) that people are mistaken in their assessments of the comparative value of different experiences containing differing durations and intensities of pleasure and pain, because according to non-cognitivism there are no objective values out there in the world for people to be mistaken about, and no objective reasons for preferring one thing to another. If ethics is thus a branch of aesthetics, then claiming that Jones is mistaken about how much happiness it is rational to personally forgo to bring X amount of happiness to Smith is like claiming that Jones is mistaken in preferring to look at black velvet clown paintings rather than Picassos. Mistakes are impossible in this sort of judgment.

One sort of non-cognitivist utilitarianism might claim to avoid this objection. Suppose that one held, first, that to say that I ought to do a given thing is, as a reforming definism, to say that it is in accord with the morality that I would choose

for my society on the hypothesis that I were calm and epistemically rational, and, second, that the moral code one would choose under these hypothetical circumstances is some form of rule utilitarian short moral code. It is stipulated of the epistemically rational person that its choices will not be affected by logical errors or by failure to make the logical inferences relevant to (i.e., capable of influencing) them; it will have all the conceptual apparatus relevant to these choices; and it will have the relevant factual information as well as a vivid idea of the relevant consequences of its choice (paraphrasing Sikora 1978, p. 148-50). By this version of non-cognitivist utilitarianism, even if pleasure and pain were completely objectively measurable, some people would have more generous sacrifice/gain ratios than others. Proponents of this theory might hold that since each person has a subjectively rational sacrifice/gain ratio, a ratio they personally want to use to guide their own behavior, the utilitarian set of rules could ensure that the correct sacrifice/gain ratio is, to an acceptable approximation, achieved by everybody. The set of rules would do this by specifying some minimum sacrifice/gain ratio, low enough that almost anyone will want to achieve it (and with modest effort will be able to), as obligatory, and permitting individuals to sacrifice more if they wish by supererogation. However, this sort of non-cognitivist utilitarianism has the same flaw as so many other contemporary metaethics, namely that it is unwittingly a form of strong moral

objectivism, a cognitivism. It assumes that certain facts yield reasons for action, in that if I want x to be the case, then I have a reason for acting to bring about x. Consider the following facts:

1. Jones wants x to occur
2. I am Jones
3. That person over there is Jones

Somehow, according to the sort of utilitarianism under discussion, facts 1 and 2 jointly make rational my choice of bringing about x, in that they provide me with a reason for action, such as the act of pursuing a certain sacrifice/gain ratio. But to call something a reason for action is to suggest that it provides some consideration beyond the mere causal power of the want to generate the act, namely some sort of conceptual fittingness between want and act. If one allows conceptual fittingness as a coherent notion, then one must also admit the metaphysical adequacy of altruistic reasons for action, derived from, e.g., facts 1 and 3 above, resulting in Jones' wants giving me a reason to act to bring about the things Jones wants. These objective, interpersonal values might then turn out to yield, as OCU claims, the same sacrifice/gain ratio for everyone. If the above sort of 'non-cognitivist' utilitarianism wishes to reject its unwitting cognitivism, then it must stop calling the agent's wants reasons for action, hence must give up the claim that there are differences in individuals' sacrifice/gain ratios, for it has only shown that there are differences in what individuals want

their sacrifice/gain ratios to be, not that there are differences in what individuals have reason to make their sacrifice/gain ratios.

Thus this sort of 'non-cognitivist' utilitarianism can generate no coherent objection to OCU's claim that there is a sacrifice/gain ratio that is correct for everyone. Moreover, OCU can offer a coherent explanation why this claim is true. As a cognitivist metaethic, OCU can and does claim that when people differ in their opinions about how much happiness one should sacrifice to bring about a given quantity of happiness for someone else, at least one of the people differing must be mistaken. Humans' asking themselves what sacrifice/gain ratio would be epistemically rational for them would be like humans' asking themselves what square root of 25 would be epistemically rational for them.

Thus I claim that a non-cognitivist theory can allow interpersonal comparisons of many aspects of pleasure and pain, e.g., their duration, their intensity, and their status as qualia, but cannot allow interpersonal comparisons of their objective value (entified or not), for according to non-cognitivism they have none to compare. Thus under non-cognitivism, it is not the case that pleasure and pain contain any concrete intrinsic value entities, or are rationally objectively preferable to one another. Under non-cognitivism pleasure and pain have subjective value, and these subjective values can be interpersonally compared, but that is a different

matter. Thus on a non-cognitivist theory one can say, 'I like Picassos twice as much as you like them', but not 'Picassos are twice as good you think they are' or 'quantity x of pleasure is twice as good as you think it is'.

A moral cognitivist, on the other hand, can say, 'quantity x of pleasure is twice as good as you think it is' with the full weight of objective rationality.

I make use of the epistemic/ontic distinction here. I use some of the same preference tests that a non-cognitivist would use, though these tests must be reinterpreted to accommodate objectivist phenomenology and metaphysics. I claim that there are concrete, objective value existents (capable of grounding ontic obligations), and objective value experiences (capable of grounding epistemic obligations); value experiences involve entertainments of the concept of objective value. Objective value experiences, as opposed to aesthetic value experiences, are the raw material of tests for ontic values; though not always veridical, objective value experiences are evidence for the existence and nature of objective value entities, as discussed in ch. 3.

I leave it open that other tests of objective value might exist over and above those adaptable from non-cognitivists, but I do not at present know of any. The two sorts of test, cognitivist and non-cognitivist, are closely related in that each can be derived from the other by suitable reinterpretation, but neither has any important nonethical priority over the other -

they have both been around for a long time.

An intuitive fear about OCU's sort of solution to the problem of supererogation is that some unlucky person is going to get stuck in life circumstances that make it clear that by undergoing a vast amount of unhappiness they can create a slightly larger amount of happiness for others. If the sacrifice/gain ratio of anything less than 1 is correct, then surely OCU must tell the unlucky person that they must make the sacrifice and have a most unhappy life. However, this fear ignores the fact that no rule demanding such an enormous sacrifice from a single human would be obeyed very often, because the slight net gain in happiness would seldom be able to causally overcome the vast fear of pain that most humans do and should have as a means of detecting and minimizing bodily injury. Therefore such a rule would not long endure, hence would not be included in the set of rules endorsed by OCU for humans (or, for that matter, any other causally possible species of intelligent sentient). A chief purpose of describing in some detail the kinds of rules that would govern supererogation and altruism under OCU was to show how an appropriate set could achieve the correct ratio for society as a whole yet distribute the burden fairly evenly. No intrinsic value is attached to distributing the burden evenly - it is merely a means to getting humans to take up the burden at all.

Thus OCU can maintain that since pleasure and only pleasure has intrinsic goodness, one should seek to bring it about without

intrinsic regard to who gets it. Yet OCU can also maintain that for contingent causal reasons one should have some extrinsic regard for who gets it.

OCU does not maintain that every human should make exactly the same sorts and quantities of sacrifices. The average human should follow the basic rules as set out in the short moral code, but able, strong-willed persons are permitted as mentioned to sacrifice more, supererogatorily, and those disabled in mind or body are permitted to sacrifice less, since they often have less happiness to make sacrifices from. This does not involve private moralities; it involves a single public morality that recognizes different social roles based upon the different kinds of minds and bodies that humans have. Other relevant social roles are those of male, female, doctor, lawyer, teacher, soldier, adult, child, and so on. A given human may occupy many of these roles either serially or, in some cases, simultaneously. Minor differences in the lower-level rights and duties of occupiers of these roles have rational bases, yield similar happiness levels, and do not supersede the equal dignity and equal higher-level rights and duties of every human. It would be a lamentably rigid and impractical moral code that did not have some such flexible features.

An essential property of such a code's recognizing different social roles is that all roles are in principle open to anyone. We are all children early in life, and merely by living long enough we can later enjoy the rights of adulthood. We may not be

born male or female, but can in principle have sex-change treatments giving us access to a new sexuality. We can all go to school and learn to become doctors or lawyers, etc. Anyone can apply to join the military. Admittedly, the facts of one's mind or body may prevent one from achieving a given social role; a dunce will fail the law exams, and a paraplegic will fail the military physical exam. However, analogously to the compatibilism by which one is free to do x, though determined to do otherwise, I propose a compatibilism by which one is free to pursue any of the social roles of epistemic RU, though on occasion determined by contingent personal characteristics not to achieve a given social role.

Able, strong-willed people are permitted to decline supererogatory acts partly because it is so hard for a human, even a strong-willed one, to be sure that they are in a situation in which an uncalled-for, genuinely gainful, hence supererogatory, sacrificial act is possible. Even when able people are intelligent and well-educated enough to be rightly sure that large gains of utility would result, they may not be strong-willed enough to overcome their strong, generally useful fear of pain.

Thus the third criticism of OCU's form of supererogation, the sacrifice/gain ratio objection, fails.

Thus all three criticisms of OCU's form of supererogation have been answered, hence the fourth objection to OCU is refuted.

Chapter Six

TWO METAPHYSICAL BACKGROUND ISSUES

1. Goodness and pluralism

The fifth objection to OCU is that moral pluralism might be true, but that OCU is a form of moral monism, hence is implausible to the degree that moral pluralism is a live option.

OCU takes no position on the issue whether a plurality of kinds of thing are good. While there is no metaphysical necessity to there being only one sort of intrinsic good, no more than that there should be only one sort of color, in the cases of putative goods other than pleasures, parsimonious entity-reduction seems in order. Few reliable intuitions would be disturbed by regarding such objects, e.g., knowledge, love, freedom, and honor, as having only extrinsic value, so long as certain relevantly related other entities, pleasures, are regarded as having intrinsic value. OCU is a form of hedonism (taken as the moral doctrine that only pleasure is objectively intrinsically good, not as one of some other doctrines often confused with moral hedonism - e.g., the psychological doctrine that all desire is for pleasure, which Nowell-Smith, 1954, p. 135, calls 'hedonism'). However, it is not obvious that one

cannot be both a hedonist and a pluralist, and it would be unwise to unnecessarily commit OCU to a premise that is not obvious, particularly when many philosophers have endorsed that premise's contrary (see Goldstein 1985, p. 49). I shall therefore show in this subsection how one might be a pluralistic hedonist, thus showing that the objection that moral pluralism might be true, and is inconsistent with OCU, is unfounded.

The question whether a hedonist is a monist or a pluralist is the question whether all pleasures are pleasant equally and in fundamentally the same way, so that pleasure is a single determinable property in its own right (though also, according to OCU, at a different level in the relational hierarchy the unique determinate of the determinable goodness); or whether the entities called 'pleasures' in English are not all pleasant in the same way, but fall under two or more same-level pleasure determinables, each also a determinate of the determinable goodness. For example, Ryle claims:

that 'pleasure' can be used to signify at least two quite different things.

(1) There is the sense in which it is commonly replaced by the verbs 'enjoy' and 'like'. To say that a person has been enjoying digging is not to say that he has been both digging and doing or experiencing something else as a concomitant or effect of the digging; it is to say that he dug with

his whole heart in his task, i.e., that he dug, wanting to dig and not wanting to do anything else (or nothing) instead. . . . His digging was his pleasure. . . .

(2) There is the sense of 'pleasure' in which it is commonly replaced by such words as 'delight', 'transport', 'rapture', 'exultation', and 'joy'. These are names of moods. . . . Connected with such moods, there exist certain feelings which are commonly described as 'thrills of pleasure', 'glows of pleasure' and so forth. (1949, p. 108-09)

These two types of pleasure might be tagged, respectively, 'desire-satisfaction' and 'joy'. Many ethicists are persuaded that certain kinds of experience such as these, all loosely known as 'pleasures', are intrinsically good, and are unlikely to be dissuaded by learning that no one pleasure property underlies all these experiences; for example, Baylis 1967, p. 444. Occam's Razor favors the collapsing of two entities into one wherever possible in a theory, and many hedonists find that their intuitions are not seriously strained by the collapsing of, e.g., desire-satisfaction and joy, into one entity, pleasure. For instance, one might adopt Nowell-Smith's claim that "to desire something is to expect it to be pleasant" (1954, p. 137), and add that the satisfaction of a desire is the obtaining of the

expected pleasure when the desired object is achieved. However, some hedonists may not find this or any comparable move plausible, and yet may still view both hedonic entities, desire-satisfaction and joy, as objectively intrinsically good.

On behalf of pluralistic hedonism, it might be objected to monistic hedonism that our pleasures are far too disparate to fall under the same pleasure determinable, that we take pleasure in far too many different things, that we like the feel of far too many different experiences. This objection stems from a misapprehension of the plausible premise that pleasures are, or are relevantly analogous to, objects capable of intentionality. A paradigmatic object capable of intentionality is a sentence, which can be about some external existent ('The cat is on the mat'), or about nothing of this nature ('Wow!'); likewise a pleasure can be about some external existent ('I find gardening pleasant'), or about nothing of this nature ('I'm in a pleasant mood today'). The objects of pleasure may indeed be radically different. However, this does not entail that the pleasures themselves are radically different. Just as the nature of sentencehood is unified and independent of the nature of the entities, if any, referred to in the sentence, the feel of pleasure is the same no matter what, if anything, occasions the pleasure (that is, the feel of pleasure is the same taking pleasure in abstraction from its intensities). On this account of pleasure, it is not the case that an experience is pleasant iff you like the way it feels (which would multiply pleasures

indefinitely - digging, etc., would all be distinct pleasures); rather, pleasure is a liked feeling that is sometimes occasioned by an experience (the feeling liked being the same for any experience occasioning it). Or alternatively, pleasure is liking the way you feel, whether that feeling includes or is occasioned by an experience or not (the liking is the unchanging pleasure).

Goldstein 1985, p. 52, claims that various kinds of pleasure might be radically different but nonetheless encompassed by the same defining account. I agree, but this claim does not entail that pleasures are similar enough to fall under the same pleasure determinable. I am inclined to think that they are similar enough, that the pleasures have something profoundly in common, hence that monistic hedonism is more plausible than pluralistic hedonism.

If a unified account of pleasure cannot be made to work, then, if one wishes to remain a hedonist, one must be a pluralistic hedonist. How might hedonistic pluralism be objected to?

It might be objected that if two or more independent hedonic entities are each objectively intrinsically good, then it will sometimes be impossible to rationally calculate what to do, since some of the several hedonic goods will be incommensurate. This impossibility would be inconsistent with OCU and many other forms of utilitarianism, according to which it is always possible in principle to rationally calculate what to do. This objection takes several variations depending how incommensurability is

defined.

One definition of incommensurability is given by Raz: "A and B are incommensurate if it is neither true that one is better than the other nor that they are of equal value" (1986, p. 117). Such 'ordinal incommensurability' (adapting Hall's term - 1967, p. 38) does permit rational action of a weak sort: "rational action is action for (what the agent takes to be) an undefeated reason. It is not necessarily action for a reason which defeats all others" (Raz 1986, p. 132). However, ordinal incommensurability does preclude rational choice of the strong, positive sort OCU claims to provide. OCU, like other strong moral objectivisms, claims not merely that the actions it prescribes can't be proven to be foolish choices, it claims that these actions can be proven to be wise choices, that these actions can be given reasons which do defeat all others. Even the arbitrary tie-breaker choice between options of equal highest value by adherents of OCU (or of any utilitarian version of entifying moral objectivism) is strongly, positively rational in that the agent is assured that all options other than the equivaluable set yield less value, and that within the equivaluable set the choice is merely between different tokens of the same type (goodness embedded in a pleasure); no wild card is in play, so one is rationally guaranteed that no important difference lurks between the options. Ordinally incommensurate options are unsettling precisely because we know that important differences may lurk between them, yet have no rational way of

assessing these differences.

Ordinal incommensurability arises frequently in epistemic contexts. A human, even a convinced utilitarian, may be forced to choose between a friendship and a better-paying job, yet due to limited evidence have no way of rationally preferring one option to the other. This situation passes the transitivity test of incommensurability (Raz 1986, p. 120-21); although reason is indifferent between the two options, it does not rank them equals. Increasing the job's salary would make the job better than it had been, but would not by transitivity thereby make the new option superior to its predecessor's rival in the eyes of the human agent. The agent choosing simply has inadequate evidence to yield a rational preference between friendship and money.

However, epistemic ordinal incommensurability does not entail ontic ordinal incommensurability. The unavailability of evidence permitting, e.g., friendship and money to be humanly compared may be merely contingent. OCU holds that all goods are in principle ordinally commensurate, hence that an omniscient agent could choose rationally from any set of options. This line assumes that people may be mistaken in their assessments of a value; to this assumption Raz objects that,

the very assumption of a possible gap
between people's considered judgments of the
comparative value of options and their real
values presupposes that values have a
reality which is independent of people's

perceptions of their lives and the value of options which are in principle open to them. This presupposition, the Platonic assumption we can call it, is to be rejected. (1986, p. 132)

However, as shown earlier in this thesis, there is no compelling reason to reject real moral values, and much to be metaethically gained from accepting them. OCU is therefore unharmed by this objection derived from the first definition of incommensurability.

A second type of incommensurability is 'linear incommensurability' (again adapting from Hall 1967): 'A and B are incommensurate if they lack an appropriate linear dimension permitting numeric quantification of value'. This definition grounds several objections to OCU, including an objection to OCU's version of hedonic pluralism. First, it might be objected to OCU (and many other utilitarianisms) that they speak freely of numeric quantities of pleasure, e.g., in locutions like 'suppose that Jones obtains 100 hedons of pleasure from an hour's walk in the woods', implying that pleasures are commensurate in this second sense. Yet, goes the objection, pleasures lack an appropriate linear dimension by which their value might be numerically quantified, hence are incommensurate in this second sense; therefore OCU must be false. This objection arises from the fact that though we may casually say things like 'it feels twice as hot today as it did yesterday', such talk is literal

nonsense since heat sensations and their ilk have no intrinsic quantitative values. The correspondence of heat sensations to numeric temperatures such as those of the centigrade or fahrenheit scales is merely conventional; unlike thermometers, heat sensations seem to have no appropriate intrinsic linear dimension. Locutions such as 'it's twice as pleasant to watch playoff hockey as it is to watch regular-season hockey' have a similar form to such nonsense locutions, which suggests that a similar confusion may underlie such pleasure locutions, via the lack of an appropriate intrinsic linear dimension to pleasure.

For purposes of argument, let us grant that the only linear dimension of pleasure available even as a candidate for appropriateness is the temporal dimension; then we shall see whether the objection can nonetheless be answered. Under the assumption granted, every pleasure has this temporal dimension, for every pleasant event takes place in time, hence is temporally extended.

Yet hedons (= utils, etc.), interpreted as measurements of pleasure, measure more than the mere temporal length of a pleasure; we want to be able to say that Jones obtained 100 hedons from an hour's walk in the woods with Smith, but that Smith obtained 200 from the same walk, so that Smith's mental states were on average twice as pleasant as Jones's during each time-slice of the walk. The extra thing being measured is the intensity of the pleasure; twice as intense is twice as good. Hedons are numerically quantitative units, so presuppose a linear

dimension (a number line) in whatever they measure, yet by hypothesis pleasure has no potentially appropriate intrinsic linear dimension beyond time. Since the extra thing being measured by hedons therefore (goes the objection) both is and is not an intrinsic linear dimension, the concept of a hedon collapses in self-contradiction. And without hedons, OCU lacks a way of comparing the quantities of different pleasures.

To help answer this objection from linear incommensurability, let us examine in its light a popular utilitarian method of accomplishing the numeric quantification of pleasure (e.g., Hall 1967, p. 44-45). According to this method, one pleasant mental state is correctly measured as containing, e.g., twice the amount of pleasure per time-unit as another pleasant mental state if, all else equal, experienced agents are indifferent between obtaining one time-unit of the former pleasant mental state and two time-units of the latter. The combination of this way of quantifying the intensity of pleasure and the normal quantification of time-segments allows all pleasures to be quantitatively compared; thus hedonistic utilitarians can make rational choices between options.

OCU's version of this sample method of comparison might be criticized on the ground that it is liable to the error of being a form of ideal observer theory, in that it tries to substitute, for an objective intrinsic consideration, the subjective response, to the relevant entities, of some type of observer. OCU responds that while this comparison method must, to be

acceptable, show that the measures produced are correct because they correspond to some unstated underlying objective consideration between the relevant entities, this correspondence arguably obtains. However, positing such an underlying ontic feature of pleasure exposes OCU's version of this comparison method to the objection just described that pleasures are linearly incommensurate.

To escape this objection, OCU must hold either that (1) hedons despite their superficial numeric character do not measure numeric quantities, or (2) that numeric quantities may measure something other than a linear dimension. To see which premise OCU should embrace, let's examine what OCU requires from hedons in pursuit of rational choice.

What OCU requires is a rational value ordering of the members of the power set of all possible tokens of simple hedonic items (hence including repetitions of some such items), so that in any possible ethical choice between two sets of consequences, one set will be objectively more valuable than the other, or the two will be equally valuable. Since the set of all possible tokens of simple hedonic items may be infinitely large, this expression of OCU's requirements is logically problematic, for standard logic applies only to finite systems. But let's assume that this problem can be overcome; it is common to all consequentialisms, hence can be overcome if, as they do, consequences make a practical moral difference. Moreover, logic applies to the universe of physical objects, which may well be

infinitely large; if we do discover it to be infinitely large, then physical science will be liable to the same objection, and surely we are not prepared to give up physical science just because of the discovery of a lot more stuff pretty much the same as the stuff physical science already explains so well.

Given such a value ordering, moral agents may rationally choose a course of action in any possible moral dilemma. Hedons function to express such a value ordering in such a way that adding and subtracting numbers of hedons, in correspondence with the addition or subtraction from an option of the hedonic events assigned these hedonic numbers, preserves the accuracy of the value ordering. For example, assume the following options and value orderings (event '2X' is 2 tokens of event X; event '0' is a mental event that is neither pleasant nor painful):

option 1 = events A, B, and C

option 2 = events A, B and D

option 3 = events B and D

option 4 = events A, B, and 2C

$$0 < A < B = 2A < C = 3B < D = 2C$$

Assume that each of these events is one time-unit's worth of the relevant event type. Assume also that these events do not change values when in combination with other events. This assumption still allows the possibility of valuable organic wholes; however,

any such organic wholes would be listed separately. Thus, e.g., the value of option 3 is captured entirely in the values of events B and D, so that the 'event' B&D has no value (if it did, option 3 would be described as events B, D, and B&D). Option 2 is thus of greater value than options 1 or 3, and of the same value as option 4. What OCU requires is an assignment of hedonic value numbers to events A-D such that for any two options w and z, the sum of the hedonic value numbers of w's events is less than the sum of the hedonic value numbers of z's events iff $w < z$; likewise mutatis mutandis for ' $w = z$ ' and ' $w > z$ '. But given the assumptions made about value orders logically 'behaving themselves' via association, transitivity and the like, it is easy to construct such an assignment. Simply assign some hedonic event X of 1 (arbitrary) time-unit's duration the value of 1 hedon, and assign hedonic value numbers to all other hedonic event types in inverse proportion to the length of the time-segment they need to equal X in value. This assignment yields the comparative worth of a standard temporal length of any hedonic event type, in such a way as to preserve the value orderings of all possible options through all possible additions or subtractions of events and their associated hedonic numbers. And as explained in reply to the objection from ordinal incommensurability, the mild logical ill behavior of humanly perceived values is, in the context of hedonism, an epistemic phenomenon, consistent with complete logical well-behavedness in the underlying ontic phenomena of objective values; moreover,

this assumption of complete logical well-behavedness is rational (see Hall 1967, p. 45-51). OCU thus gets what it needs from hedonic numbers without assuming that pleasures have a second linear dimension of the sort objected to.

In its escape from the objection from linear incommensurability, OCU thus holds that regarding the comparative worths of hedonic events of the same temporal length (in effect, the comparative worths of pleasure intensities), hedons despite their superficial numeric character do not measure numeric quantities. Instead, they use numbers to express orderings. These order expressions are true because they correspond to real moral value orderings in pleasures. The full numeric potential of hedons is used only regarding the temporal dimension. This is not to say that, regarding intensity, hedons do not measure quantities, merely orders; as C. Stumpf says, "increase and diminution are names for quantitative changes" (see Husserl 1970, p. 440). Since intensities may increase and diminish, intensities are quantitative entities. It seems natural to express this as intensities having non-numeric quantities, and line segments having numeric quantities. However, one might alternatively express the distinction as intensities' quantities being merely ordinally numeric, and line segments' quantities being linearly numeric; either way, there is a clear boundary which makes no trouble for hedons. Note that if the boundary did make trouble for hedons, it would also make trouble for some standard theoretical entities from physical science, such as

measures of heat-energy (Hall 1967, p. 40).

Thus so long as quantitative attributions of value to hedonic consequences are interpreted merely as expressions of value equi-ordering between certain temporal quantities of various hedonic items, these quantitative attributions will not entail that the valuable hedonic items have, in addition to an intrinsic value ordering, an intrinsic linear dimension present even in a time-slice of the hedonic item, as is the case in a time-slice of a yardstick. OCU (and many other utilitarianisms) may therefore avoid the objection that hedonic items have no intrinsic linear dimension. Confirming evidence comes from the fact that in practice we have no difficulty making at least some sorts of quantitative comparisons between hedonic consequences; e.g., torturing a person to death is at least twice as bad as stepping on their toe. Exact quantitative comparisons are often difficult or impossible in practice, but this problem is common to all the quantitative sciences.

A final objection to this method of quantitatively comparing pleasures is that if, as deemed possible earlier, pleasures are encompassed not by a single property of pleasure but by a plurality of properties such as desire-satisfaction and joy, hedons could not possibly in any way measure amounts or intensities of pleasure, because there is no unified pleasure entity to be measured. OCU replies that what hedons rank are not, strictly speaking, pleasures, but their goodnesses. Goodness is instantiated in time, so may be temporally

quantitatively measured, and is instantiated in various intensities, so may be ordinally ranked as explained above. And goodness is part of each hedonic entity whether or not these pleasures fall under a single pleasure property. For each kind of pleasure, one or many, a certain intensity of pleasure is necessarily accompanied by a certain intensity of goodness; as the intensity of the pleasure increases, so proportionately does the intensity of the goodness. As explained in ch. 2, OCU leaves it an open question whether the goodness of a pleasure is found as a part within a whole, or merely by entailment supervening on the pleasure; either way, any indexical location in which pleasure is instantiated necessarily also instantiates goodness. And these goods are what hedons measure; it is always rational to choose the most goodness available - the most temporal length and the most intensity. This system of measurement would work even if other mental items were good besides hedonic items, though I cannot imagine what such things might be (in agreement with Blake 1967, p. 437). Since, as Sidgwick notes, "it is reasonable to prefer pleasures in proportion to their intensity" (1874, p. 127), hedonic judgments can be made by agents cognizant of their awarenesses of pleasure but not of goodness, as is sometimes the case. The form of hedonic pluralism thus produced would be benign because it retains the traditional strength of hedonic monism: "one of the main attractions of traditional hedonism was its attempt to reduce all value to a single source" (Goldstein 1985, p. 50), while allowing, as some find phenomenologically

necessary, a plurality of pleasure properties.

A third definition of incommensurability is used in mathematics: 'A and B are incommensurate quantities if their ratio cannot be expressed as an ordinary fraction'. Any rational number is incommensurate in this third sense with any irrational number (call this 'mathematic incommensurability'); for instance, the numbers 1 and root 2 are mathematically incommensurate. However, mathematic incommensurability poses no significant problem for any type of utilitarianism, for it is not a claim that the relevant quantities do not exist, merely the claim that they cannot be expressed in a certain way. The numbers 1 and root 2 pick out certain points in a linear dimension that undeniably exist, hence measure line segment lengths (Meserve et al 1989, p. 114). These line segments exist and form certain proportions of each other whether or not these proportions are expressible as ordinary fractions. We can easily use other symbols to name such proportions (Meserve et al 1989, p. 125). Likewise, claims utilitarianism, certain temporal segments of intrinsic goods exist and form certain proportions of each other whether or not these proportions are expressible as ordinary fractions. There is even less difficulty for hedonic quantities expressing ranked intensities, for no underlying numeric quantities are posited.

A fourth definition of incommensurability is given by Kenshur:

to prove incommensurability, one has,

ultimately, to prove the nonexistence of any standpoint outside individual theories or conceptual systems from which they can be measured against one another. (1984, p. 376)

Such incommensurability (call it 'theoretic incommensurability') could be asserted of the moral theories that desire-satisfaction is good and that joy is good (and so forth for whatever goods the pluralist posits). However, it need not be asserted; the pluralist might be a rigorous rational objectivist about basic epistemology and metaphysics, and claim only to have distinguished between two (or more) sorts of mental state easily comprehended within one overall world-view. If, on the other hand, the pluralist opts for theoretic incommensurability regarding moral goods, they are committed within that sphere to a deep Feyerabendian relativism that appears to contradict itself (Kenshur 1984, p. 378), being yet another incarnation of naive Cartesian skepticism, the view that anything that can be doubted should be doubted. One might attempt an intermediate position according to which an overall world-view is fine, but not an overall moral view, so that every moral agent's assignment of value is equally well justified in the weak, negative sense that no justification is possible. But such an intermediate view amounts to the reduction of ethics to a branch of aesthetics, which view was rejected in §5.3 as inconsistent with moral cognitivism, and is false if even the weakest form of strong moral objectivism, e.g., Mackiean non-entifying moral

objectivism, rational prudence, is true.

Another form of incommensurability is embodied in the view that goodness is not the only fundamental moral value. For instance, one might hold that goodness and rightness are both fundamental moral values, that goodness is a consequentialist consideration, that rightness is a deontic consideration, and that there is no rational way of settling the conflicts between the two that arise in normal human life. Perhaps this form of incommensurability is assimilable to one of the previous four types, perhaps not; in either case, no such problem arises for OCU or any similar utilitarian morality, for in such moralities rightness is always derived from goodness, as explained for OCU in §1.2 (Moore 1903, p. 25, 147, agrees).

A general form of incommensurability objection has been made to utilitarianisms both hedonic and non-hedonic, e.g., by Taylor (1989, p. 230), who claims that things such as integrity, charity, and liberation are worthy in ways some of which are incommensurate, hence jointly uncapturable in the utilitarian calculus. Again, perhaps this form of incommensurability is assimilable to one of the four types discussed, perhaps not; in either case, no such problem arises for OCU or any other hedonic utilitarianism. Hedonic utilitarianism answers this objection by holding, first, that such things as integrity, charity, and liberation have no intrinsic value beyond their hedonic content, if any, and that most or all of their value is extrinsic, realized in their hedonic consequences, and second, that no

hedonic values are incommensurate. Also, as noted by Blake (1967, p. 434-35), the general objection applies to any form of consequentialism, and it is surely false that consequences are irrelevant to practical moral thinking, as they would be if incommensurability prevented us from telling better from worse acts by their consequences.

Thus the pluralistic hedonist who subscribes to OCU has nothing to fear from the various objections to value commensurability. As a result, OCU has nothing to fear from pluralistic hedonism. Since OCU is consistent with (at least one form of) moral pluralism, the fifth objection is refuted.

2. The Synthetic A Priori

The sixth objection to OCU is that OCU assumes the existence of some synthetic a priori propositions, but that no such propositions exist. OCU does make this assumption, so must defend the synthetic a priori. OCU makes this assumption by assuming that determinable properties exist; it will be seen that determinable properties, whether construed as supervening on their determinates or as existing within them as parts, stand in metaphysically necessary relations with their determinates, or with those parts of their determinates that are ontologically independent, and that the propositions describing these states of affairs are synthetic a priori. OCU also assumes the existence

of analytic a priori propositions, but these are much less controversial entities, so will not be defended in this thesis.

The question whether any propositions are synthetic a priori has been answered positively and plausibly by Sikora (1981), but let me provide a brief argument for this positive answer, incorporating, but not restricted to, some of his points. A proposition was characterized in §2.1 as analytic if its logical form, as revealed by the meanings or meaningful uses of its terms, reveals its truth-value, and synthetic otherwise. I take a priori knowledge of a proposition's truth-value to rest essentially on a correct understanding of the meanings or meaningful uses of the proposition's terms. Any proposition whose truth-value is knowable in this way, i.e., by reason unenriched by experience beyond the bare understanding of the proposition, is a priori; all other propositions are empirical. The concepts of analyticity, syntheticity, and a prioricity have been defined in many ways (Pap 1958, ch. 5). The reasons why the definitions given above are right for present purposes are complex, but let me say at least a few words on the subject.

First, I take a proposition to be analytic if its logical form, as revealed by the meanings or meaningful uses of its terms, reveals its truth-value, and synthetic otherwise; my definitional reference to logical form might be objected to on the ground that:

on the basis of such a definition, to say
that logical principles are analytic would be

no more informative than to say that logical principles are logical principles, and (the various) philosophers who maintained that logical truths are analytic surely intended to make a significant statement, a statement clarifying the nature of logical truth. (Pap 1958, p. 106)

However, two considerations justify the definitional reference to logical form.

First, one can maintain that it is significant that logical truths and falsehoods are analytic, and hold that the significance of the remark is that it expresses the interesting truth that two metaphysical terms from different contexts are being used to point at the same fundamental state of affairs. Likewise, it is interesting to learn that ' $x^2 + y^2 = 1$ ' and 'circle' point at the same fundamental state of affairs. This sort of discovery may allow proofs to link up hitherto separate domains of inquiry. As Sikora points out in this context (1981, p. 447), the members of some perfectly legitimate sets of terms cannot be defined via terms from outside their set. There are vicious circles and benign circles; one form of benign definitional circle is a dictionary.

Second, we may further illuminate the terms 'analytic' and 'logical' by linking them with the qualon/indexon distinction introduced in §1.3. Recall that a qualon is by definition an object composed wholly of qualitative properties, and that an

indexon is by definition an object composed wholly of indexical properties. Qualitative properties ontically express what a thing is, what kind of concrete stuff it is made of; indexical properties ontically express where, in a field of existence, possible or actual, a thing is. The logical form in virtue of which the truth-value of an analytic proposition is revealed is a form manifested as an indexon. All states of affairs manifesting indexons, and nothing but indexons, thereby manifest logical forms. Sikora makes the same point in different language when he says that

in the case of analytic truths, unlike sap
[synthetic a priori] truths, you don't need
to know what particular properties are
connoted by the various terms so long as you
know which of the various property
requirements are identical and which are not.

(1981, p. 451)

Husserl, also, links his version of the qualon/indexon distinction (see §1.3) to logic and the analytic/synthetic distinction (1970, p. 455-58). Such links may be discerned less explicitly in the writings of many philosophers; e.g., Hare: 'No amount of logic will show us the difference between blue and red' (1981, p. 3).

My second comment on why I use the definitions given of analyticity, syntheticity, and a prioricity is that it is better to make these definitions accommodate both the truth-values,

truth and falsity, rather than just one. I take a proposition to be analytic if its truth-value is suitably revealed to the inquiring mind, whether the truth-value revealed is truth or falsity, whereas some define analyticity in such a way that only true propositions are eligible (e.g., Bennett 1959, p. 172; Lewis 1946, p. 35). My feeling here is that while one may say some interesting things on the subject of analyticity while restricting the scope of one's remarks to truths, there is no obvious reason to so restrict one's remarks, for the points of interest are not about truth itself, but about some ways in which truth is manifested and revealed. These ways apply equally, *mutatis mutandis*, to truth and falsity, hence are better stated in the more general form, so as to better focus our attention on the relevant phenomenon. Also, it would be impossible to define, e.g., 'a priori' in a way that usefully distinguished between truths and falsehoods according to how we learn them. If, for example, we learn that it is true that ' $\neg(p.\neg p)$ ' by reason alone, then surely we may also learn that it is false that ' $p.\neg p$ ' by reason alone. Since each such truth learned generates a falsehood learned, and vice versa, most of our comments on the a priori may for brevity be phrased in terms of only one truth-value. However, such comments strictly speaking apply to both truth-values; likewise, *mutatis mutandis*, for the terms 'empirical', 'analytic', and 'synthetic'.

My third comment on why I use the definitions given of analyticity, syntheticity, and a prioricity concerns the third of

these items. I take a priori knowledge of a proposition's truth-value to rest essentially on a correct understanding of the meanings or meaningful uses of the proposition's terms. Any proposition whose truth-value is knowable in this way, i.e., by reason unenriched by experience beyond the bare understanding of the proposition, is a priori; all other propositions are empirical. One of my reasons for using this definition is that, as Sikora says (1981, p. 444), the notion of a prioricity is an epistemological notion about how we know certain things. My definition gives an account of the process by which we come to know the relevant things. Another way of putting it would be to define, as Sikora does (1981, p. 445), a priori knowledge as knowledge which lacks or does not require any inductive support. While it is true that a priori knowledge lacks or does not require any inductive support, it, as knowledge, does require some kind of support. A definition in the present context may helpfully give a positive account of the kind of support a priori knowledge rests on; my definition tries to give such an account. Sikora's definition accurately locates the boundary between a priori and empirical knowledge, and gives a positive account of what is on the empirical side of the boundary, but says little about what is on the a priori side. For some purposes this is fine - it may be sufficient to locate the boundary accurately, which Sikora's definition does concisely, and have a positive account of what is on the empirical side of the boundary. But for other purposes, such as those of my argument of §2.3

concerning the third account of resemblance, it helps to also give a positive account of the kind of support a priori knowledge rests on. Part of the purpose of this section is to see that each side of the a priori/empirical boundary is given a positive account. Similarly, my definition of 'synthetic' was negative; a positive account will be given later in this chapter.

My fourth comment on why I use the definitions given of analyticity, syntheticity, and a prioricity also concerns the third of these items. The a priori/empirical distinction is sometimes restricted in scope to propositions whose truth-values are humanly knowable, so that propositions whose truth-values are not humanly knowable are neither a priori nor empirical (e.g., Bradley and Swartz 1979, p. 150). My definitions avoid this restriction because the nonexhaustiveness of the distinction so construed is unnecessary and moves one's attention away from the epistemic properties of the proposition, where attention should be, and toward the epistemic properties of the potential knower, where it should not be. Attention should be on the epistemic properties of the proposition in the sense that in order to appreciate the full epistemic potential of propositions, we must consider their interactions with the widest possible selection of types of agents seeking knowledge. This precludes our considering only the human type. As philosophers we should seek to appreciate the full epistemic potential of propositions, and the terms 'a priori' and 'empirical', if intended to be about propositions' amenability to rational inquiry, should thus

consider propositions' interactions with all types of rational agent, including the most fully, hence superhumanly, rational. I take 'a priori' and 'empirical' to be terms of metaepistemology, whereas questions about the limits of human rationality are addressed in normative epistemology.

There is no metaepistemologically useful place to draw the line between what is humanly knowable and what is not. If what is humanly knowable is limited to what humans know, then the knowable/known distinction collapses and 'humanly knowable' loses its utility. If what is humanly knowable includes propositions unknown to humans, then it must, to be metaepistemologically general, include any proposition it is possible for a human to know. Since it is logically and metaphysically possible for a human to have any rational power, this construal collapses 'humanly knowable' and 'rationally knowable', again removing our reason for coining the former term. It is possible for a human to have any rational power in the sense that whether one is considering logical possibility, causal possibility, or whatever scope of possibility one pleases, any rational power within that scope can be combined with human possession of that power. Humans can change; we can, if we wish, add devices for physical abilities we lack, for senses we lack, or for rational powers we lack. There is no obvious inconsistency between the concepts of an agent's being human and an agent's possessing a given rational power. We already count, as humanly knowable, propositions requiring the epistemic cooperation of devices such as

binoculars, electron microscopes, and computers (e.g., Bradley and Swartz 1979, p. 149); how could we consistently admit these devices but not a device yielding some other epistemically useful power, such as a given rational power?

Assuming these accounts of the analytic/synthetic and a priori/empirical distinctions, what examples might be given of synthetic a priori propositions? They would be propositions whose truth-values are knowable by reason unenriched by experience beyond the bare understanding of the proposition, but are not revealed by its logical form. A candidate synthetic a priori proposition is:

(1) Every hue has an intensity.

One may know what a color hue is and what an intensity is, and know that the logical form of (1) does not reveal its truth-value, yet also see that anything with a hue must also have copresent, intermingled, and, in Husserl's term, 'interpenetrating' (1970, p. 437), an intensity, hence see that (1) is true.

It might be objected to the claim that such propositions as (1) are synthetic that they are expressions from a natural language whose words have meanings and meaningful uses that are in their details poorly understood, evolved for social rather than purely philosophic functions, and perhaps have or lack analytic entailments in such fashion as to render my claims about such propositions false, even though such entailments or their lacks are distracting linguistic baggages which might be

discarded to the gain of conceptual precision - that is, to the gain of clarity and conciseness, and without begging the question in favor of the synthetic a priori. I agree that this is possible, hence stipulate, in agreement with Sikora, that the propositions in question are to be regarded as expressed in an artificial language purged of such distracting analytic entailments or their lacks. For instance, in the case of proposition (1), saying in ordinary English that something has a given 'hue' may well logically imply that it has some intensity; this requirement is explicitly rejected in the artificial language proposed. The linguistic rule of usage discarded, we can then examine the property of hue - the core property ostensibly required for the satisfactory application of 'hue' - and see whether it contains as a part the property of intensity, and if not, whether it must nonetheless by metaphysical necessity be accompanied by the property of intensity. Sikora uses hue and intensity to explain how core properties may be identified and distracting linguistic rules of usage discarded:

You are shown two batches of color swatches.

All the swatches in the first group have the same determinate hue but there is no other determinate quality common to all of them.

In the second batch they all have the same determinate intensity but again this is the only determinate quality common to all of them. You are told that 'L-red-156' is to

stand for the common determinate quality of the first batch and that 'L-intensity' is to stand for the determinate quality of which the determinate quality common to all members of the second batch is a determinate version. Further, to avoid ambiguity, you are told that although all the instantiations of L-red-156 that you have been shown have some L-intensity, saying that something is L-red-156 is not to be taken as logically implying that it has some L-intensity. Saying in ordinary English that something has a given hue may well logically imply that it has some intensity, but this requirement is explicitly rejected for the artificial terminology in question. Given these stipulations it would not be contradictory to say that something could be L-red-156 yet lacked an L-intensity. None of the rules governing the use of the two key terms would preclude such an object: you didn't stipulate that for something to be L-red-156 it must have an L-intensity, only that it must have the determinate quality common to the first batch of swatches. . . . Thus it is a synthetic truth that anything that is

L-red-156 must have an L-intensity, and since it is not an empirical truth it is a synthetic a priori truth. (1981, p. 450-51)

Thus sloppy color words from ordinary English, such as 'hue' and 'intensity', may be used to help define theoretical terms precise enough to permit inferences of philosophic rigor, much as was claimed in §1.2 of sloppy moral words from ordinary English.

It might be objected to synthetic a priori claims, e.g., of the form of (1), asserting that one thing must always be accompanied by another thing, that the necessity is an artifact of consciousness, that the mere fact that we cannot, e.g., imagine the one thing without the other, demonstrates a limit of our imagination, not of external reality. But as Husserl says,

 differences such as this, that one object
 . . . can be 'in and for itself', while
 another can only have being in, or attached
 to some other object - are no mere
 contingencies of our subjective thinking.
 They are real differences, grounded in the
 pure essences of things, which, since they
 obtain, and since we know of them, prompt us
 to say that a thought which oversteps them is
 impossible . . . (1970, p. 445)

As explained in §3.1, Occam's Razor presumes against the existence of all intuitions, but once granted their existence presumes in favor of their veridicality. Thus if we have the

intuition that a thing is impossible, then we should try to include in our world-view the hypothesis that the thing is impossible, not the more complicated hypothesis that the thing is possible but that to have an intuition that the thing is possible is impossible or at least counterfactual. Unless independent evidence indicates otherwise, we should accept our intuitions (assuming we have them) that some propositions, e.g., of the form of (1), are synthetic a priori. It makes no more sense to doubt these intuitions than to doubt our intuitions that analytically, necessarily $2 + 2 = 4$. My arguments regarding, e.g., (1), are designed to arouse the former intuitions in rational persons, and eliminate contrary intuitions.

It should also be noted that the necessity attributed to propositions known a priori is not epistemic, but ontic, a logical or metaphysical state of affairs an a priori proposition expresses. One can no more be certain of a priori knowledge than of empirical knowledge - a mad scientist or Cartesian demon might be tricking us in either case by supplying misleading evidence. A priori knowledge is necessary only in that the proposition in question is claimed to have the same truth-value in all possible worlds. It is irrelevant to point out that foolish people often claim to be absolutely certain of various propositions; their certainty is unwarranted by the evidence available to them. No amount of evidence warrants absolute certainty about anything, and in any case, certainty is a psychological state of a believer, not a logical or metaphysical state of affairs a

believed proposition expresses (see Pap 1958, p. 125-27).

Note that one can consistently hold that while neither a priori nor empirical knowledge can yield justified absolute certainty, some a priori knowledge can justifiably yield a psychological state nearer to absolute certainty than can any empirical knowledge. Broad, for instance, seems to hold this view; he asserts, of a way of acting, that "we know as surely as we can know anything that is not a priori, that by no means everybody will act in this way" (1916, p. 377-8). Thus one can in principle know two propositions beyond reasonable doubt yet have less doubt about one than the other. OCU allows that a priori and empirical knowledge yield different maximum degrees of human confidence, but does not insist upon this view.

The reason OCU assumes the existence of some synthetic a priori propositions is that whether a determinable accompanies its determinates as a part or as a supervening ontologically independent property, the situation contains some metaphysical necessity, expressible in a synthetic a priori proposition. It was explained in ch. 2 how the latter state of affairs, with its supervening ontologically independent determinable, is described by a synthetic a priori proposition; now consider how the former state of affairs, in which a determinate is logically necessarily accompanied by its determinable as a part, has an aspect described by a synthetic a priori proposition. Recall that on any such account, determinate-determinable propositions are analytic; e.g., the proposition 'All red things are colored' is

of the form $x((Fx.Cx) \rightarrow Cx)$ where 'C' attributes coloredness and 'F' attributes some property F which combined with coloredness makes redness. C and F are thus ontologically independent properties, and yet it is inconceivable that anything F should not also be C. Therefore $x(Fx \rightarrow Cx)$ is a synthetic a priori truth.

The only way to deny the existence of such synthetic a priori propositions is to deny either that determinate properties resemble each other in the way in question, or to deny that the determinable-property theory is the correct account of such resemblances.

The first disjunct is false. Determinate properties do resemble each other in various ways. In particular, every color property resembles every other color property in being a color. As mentioned in §2.3, Armstrong is right that "these recalcitrant resemblances seem objective phenomena, demanding an ontological analysis" (1978, vol. 2, p. 116).

The second disjunct, that these resemblances exist but have correct accounts other than variants of the determinable-property theory, is dubious because both the chief rival accounts are dubious. To show this I reviewed in §2.3 the five chief ontological accounts of resemblance. Three turned out to posit determinables, and the other two turned out to be flawed.

Thus whether a determinable accompanies its determinates as a part or as a supervening ontologically independent property, the situation contains some metaphysical necessity, expressible in a synthetic a priori proposition. While Moore did not make

use of determinable-determinate property theory in his account of goodness and goods, he did grasp that the relations between goodness and goods always contain some metaphysical necessity: "propositions about the good are all of them synthetic and never analytic" (1903, p. 7). While this claim is false because too general (consider the proposition 'all good things are good'), we are now able to reconstruct the true claim Moore was trying to make.

Finally, let us examine the alternative to the existence of synthetic a priori propositions. Since the analytic/synthetic distinction was taken to be exhaustive of propositions, the following proposition would be true:

(2) All a priori propositions are analytic.

Moreover, (2) would be knowable a priori, according to at least some skeptics of synthetic a priori propositions, e.g., M. Schlick (see Pap 1958, p. 94-95). Therefore (2) itself would be analytic (leaving aside the problem of self-referential propositions). Since analytic propositions are those whose truth-values are revealed by their logical form, and by hypothesis (2) is analytic, analyticity would therefore have to be logically contained in a prioricity. A prioricity obtains of any proposition whose truth-value is revealed by the proposition's terms' meanings or meaningful uses; analyticity obtains of any proposition whose truth-value is revealed by the proposition's logical form. A proposition is an attribution of properties to individuals. The logical form of a proposition

does not depend on the nature of the properties attributed. The nature of a property is expressed in the meanings or meaningful uses of terms. Thus the logical form of a proposition does not depend on what is expressed in some meanings or meaningful uses of terms of propositions. Hence such meanings or meaningful uses could be changed without changing the proposition's logical form. Therefore logical form is not logically contained in some meanings or meaningful uses of terms. Thus analyticity is not logically contained in a prioricity, hence (2) is not, as was hypothesized, analytic.

To put the point slightly differently, a prioricity obtains of any proposition whose truth-value is revealed by the proposition's asserted tautological indexons, contradictory indexons, metaphysically necessary qualons, or metaphysically impossible qualons; analyticity obtains of any proposition whose truth-value is revealed by the proposition's asserted tautological indexons or contradictory indexons. Since a disjunct is not logically contained in a disjunction in which it appears (that is, p or q does not logically imply q), a prioricity does not logically contain analyticity, hence (2) is not, as was hypothesized, analytic.

One might try to fall back on the position that (2) is true but empirical rather than a priori, but this position is difficult to state coherently.

Empirical propositions are usually thought of as being true in some possible worlds and false in some possible worlds,

experiential evidence being needed to determine which of these two sets of worlds the actual world belongs to. But if (2) is false in some possible world, then in some world some a priori proposition is not analytic, i.e., is synthetic. But if a given proposition is synthetic a priori in some possible worlds, surely it is synthetic a priori in all possible worlds in which it exists, for synthetic a prioricity is a way in which a proposition is via its intrinsic nature knowable. One might hold that the relevant proposition just happens not to exist in the actual world, but then one must reject the widely-accepted premise that any proposition that exists in one possible world exists in all possible worlds (e.g., Bradley and Swartz 1979, p. 13-24).

Alternatively, one might characterize empirical propositions as including propositions which are true in all possible worlds, but nonetheless in need of experiential evidence for assessment of their truth-values. But if it is impossible for such propositions to be false, why would this impossibility not be revealed in their terms' meanings or meaningful uses, hence knowable a priori, hence not in need of experiential evidence for assessment of their truth-values? Analytic impossibility would be so revealed, as would synthetic impossibility; is there some other relevant kind of impossibility? A coherent account of such impossibility would be needed before this position could be accepted.

Thus the sixth objection to OCU is refuted.

3. Conclusion

To conclude this thesis, OCU is a viable comprehensive moral theory, able to answer the six objections currently most influential against it among metaethicists and some others. It was also revealed in the course of this investigation that many alleged competitors to OCU for the status of a plausible comprehensive moral theory are not in fact comprehensive, but duck or beg central metaethical questions. The remaining competitors to OCU suffer important unanswered objections; therefore OCU can justifiably claim to be the most plausible comprehensive moral theory.

BIBLIOGRAPHY

- Angel, Leonard (1989), How to Build a Conscious Machine.
Boulder: Westview.
- Aristotle (1985), Nicomachean Ethics. Trans. Terence Irwin.
Indianapolis: Hackett.
- Armstrong, David M. (1978), Universals and Scientific Realism.
Vol. 2. Cambridge: Cambridge UP. 2 vols.
———. (1989), Universals: An Opinionated Introduction.
Boulder: Westview.
- Baier, Kurt E.M. (1958), The Moral Point of View. Ithaca, New
York: Cornell UP.
- Baylis, Charles A. (1967), 'Grading, Values, and Choice',
Problems of Moral Philosophy. Ed. Paul W. Taylor.
Belmont, California: Dickenson. 439-51.
- Bennett, Jonathan (1959), 'Analytic-Synthetic', Aristotelian
Society Proceedings 59: 163-88.
———. (1984), A Study of Spinoza's Ethics. Indianapolis:
Hackett.
- Blake, Ralph Mason (1967), 'Why Not Hedonism?', Problems of Moral
Philosophy. Ed. Paul W. Taylor. Belmont, California:
Dickenson. 431-39.
- Boyd, Neil (1991), High Society. Toronto: Seal, 1993.
- Bradley, Raymond, and Norman Swartz (1979), Possible Worlds.
Indianapolis: Hackett.
- Brandt, Richard B. (1959), Ethical Theory. Englewood Cliffs,
N.J.: Prentice.
———. (1963), 'Toward A Credible Form Of Utilitarianism',
Morality and the Language of Conduct. Eds. H.N. Castaneda
and G. Nakhnikian. Detroit: Wayne State UP. 107-43.
———. (1979), A Theory of the Good and the Right. Oxford:
Clarendon Press.
- Brentano, Franz (1889), The Origin of our Knowledge of Right and
Wrong. 3rd ed. Eds. Oskar Kraus and Roderick M. Chisholm,
Trans. Roderick M. Chisholm and Elizabeth H. Schneewind.
London: Routledge and Kegan Paul, 1969.
- Broad, C.D. (1916), 'On the Function of False Hypotheses in
Ethics', International Journal of Ethics 26: 377-97.

- Chisholm, Roderick M. (1969), 'The Loose and Popular and the Strict and Philosophical Senses of Identity', Perception and Personal Identity: Proceedings of the 1967 Oberlin Colloquium in Philosophy. Eds. Norman S. Care and Robert H. Grimm. Cleveland: Case Western Reserve UP. 82-106.
- Dancy, Jonathan (1988), 'Supererogation and Moral Realism', Human Agency. Eds. Jonathan Dancy, J.M.E. Moravcsik, and C.C.W. Taylor. Stanford, California: Stanford UP. 170-88.
- . (1991), 'Intuitionism', A Companion to Ethics. Ed. Peter Singer. Oxford: Basil Blackwell. 411-19.
- Danto, Arthur (1963), 'What We Can Do', The Journal of Philosophy 60: 435-45.
- Dennett, Daniel C. (1982), 'Mechanism and Responsibility', Free Will. Ed. Gary Watson. Oxford: Oxford UP. 150-73.
- Diggs, B.J. (1978), 'Rules and Utilitarianism', Contemporary Utilitarianism. Ed. Michael D. Bayles. Gloucester, Mass.: Peter Smith. 203-38.
- Dolby, R.G.A. (1973), 'Philosophy and the Incompatibility of Colours', Analysis 34: 8-16.
- Foley, Richard (1978), 'Prudence and the Desire Theory of Reasons', Journal of Value Inquiry 12: 68-73.
- Gewirth, Alan (1978), Reason and Morality. Chicago: U of Chicago P.
- Goldman, Alvin I. (1970), A Theory of Human Action. Englewood Cliffs, New Jersey: Prentice.
- Goldstein, Irwin (1985), 'Hedonic Pluralism', Philosophical Studies 48: 49-55.
- . (1989), 'Pleasure and Pain: Unconditional, Intrinsic Values', Philosophy and Phenomenological Research 50: 255-76.
- Gould, S.J. (1993), 'The Sexual Politics of Classification', Natural History 102,11: 20-29.
- Hall, John C. (1967), 'Quantity of Pleasure', Aristotelian Society Proceedings 67: 35-52.
- Hardin, Garrett (1985), Filters Against Folly. London: Penguin, 1987.
- Hare, R.M. (1981), Moral Thinking. Oxford: Clarendon Press.

- Harrod, R.F. (1936), 'Utilitarianism Revised', Mind 45: 137-56.
- Heyd, David (1982), Supererogation: Its Status in Ethical Theory. Cambridge: Cambridge UP.
- Hobart, R.E. (1934), 'Free will as involving determination and inconceivable without it.' Mind 43: 1-27.
- Hume, David (1888), A Treatise of Human Nature. Ed. L.A. Selby-Bigge. 2nd ed. Oxford: Clarendon Press, 1987.
- Husserl, Edmund (1970), Logical Investigations. Vol. 2. Trans. J.N. Findlay. London: Routledge & Kegan Paul. 2 vols.
- Johnson, W.E. (1921), Logic, Part I. Cambridge, U.K.: Cambridge UP.
- Kagan, Shelly (1984), 'Does Consequentialism Demand Too Much?', Philosophy and Public Affairs 13: 239-54.
- Kant, Immanuel (1948), Groundwork of the Metaphysic of Morals. Trans. H.J. Paton. New York: Harper, 1964.
- Kenshur, Oscar (1984), 'The Rhetoric of Incommensurability', Journal of Aesthetics and Art Criticism 42: 375-82.
- Kim, Jaegwon (1978), 'Nomological Incommensurables', American Philosophical Quarterly 15.2: 149-156.
- Langford, Michael J. (1988), 'Supererogation and Friendship', Philosophy and Culture 3: Proceedings of the 17th World Congress of Philosophy. Ed. V. Cauchy. 441-45.
- Lewis, Clarence Irving (1946), An Analysis of Knowledge and Valuation. La Salle, Ill.: Open Court.
- Locke, John (1894), An Essay Concerning Human Understanding. Vol. 1. Oxford: Clarendon Press. 2 vols.
- Lyons, David (1967), Forms and Limits of Utilitarianism. Oxford: Clarendon Press.
- Mackie, J.L. (1977), Ethics: Inventing Right and Wrong. Harmondsworth, Middlesex, England: Penguin.
- Mazlish, Bruce (1975), James and John Stuart Mill. New York: Basic.
- McCarty, Richard (1989), 'The Limits of Kantian Duty, and Beyond', American Philosophical Quarterly 26: 43-52.
- McDowell, John (1985), 'Values and Secondary Qualities', Morality

- and Objectivity. Ed. Ted Honderich. London: Routledge & Kegan Paul. 110-29.
- Meserve, Bruce E., Max A. Sobel, and John A. Dossey (1989), Introduction to Mathematics. 6th ed. Englewood Cliffs, New Jersey: Prentice.
- Mill, John Stuart (1861), Utilitarianism. Ed. Oskar Piest. New York: Macmillan, 1957.
- Montague, Phillip (1989), 'Acts, Agents, and Supererogation', American Philosophical Quarterly 26: 101-11.
- Moore, George Edward (1903), Principia Ethica. Cambridge: Cambridge UP, 1959.
- . (1912), Ethics. London: Oxford UP, 1963.
- . (1922), Philosophical Studies. London: Routledge & Kegan Paul, 1965.
- . (1959), 'Is Goodness a Quality?', Philosophical Papers. New York: Collier, 1966. 89-100.
- Nagel, Thomas (1970), The Possibility of Altruism. Oxford: Clarendon Press.
- . (1974), 'What Is It Like To Be a Bat?', Philosophical Review 83: 435-50.
- Neurath, Otto (1937), 'Unified Science and its Encyclopedia', Philosophy of Science 4: 265-77.
- Nowell-Smith, P.H. (1954), Ethics. Harmondsworth, Middlesex, England: Penguin, 1965.
- Oddie, Graham (1991), 'Supervenience, Goodness, and Higher-Order Universals', Australasian Journal of Philosophy 69: 20-47.
- Pap, Arthur (1958), Semantics and Necessary Truth. New Haven and London: Yale UP, 1966.
- Perry, John (1979), 'The Problem of the Essential Indexical', Nous 13: 3-21.
- Rawls, John (1971), A Theory of Justice. Cambridge, Massachusetts: Belknap-Harvard UP.
- . (1978), 'The Basic Structure as Subject', Values and Morals. Eds. A.I. Goldman, and J. Kim. Dordrecht, Holland: D. Reidel. 47-71.
- Raz, Joseph (1986), 'Value Incommensurability: Some Preliminaries', Aristotelian Society Proceedings 86: 117-34.
- Rescher, Nicholas (1973), The Coherence Theory of Truth. Oxford: Clarendon Press.

- Rorty, Richard (1989), 'Solidarity or Objectivity?', Anti-Theory in Ethics and Moral Conservatism. Eds. Stanley G. Clarke and Evan Simpson. Albany, New York: State U of New York P. 167-83.
- Ross, William David (1967), 'Prima Facie Duties', Problems of Moral Philosophy. Ed. Paul W. Taylor. Belmont, California: Dickenson. 266-78.
- Russell, Bertrand (1912), The Problems of Philosophy. New York: Oxford UP, 1966.
- . (1959), My Philosophical Development. New York: Allen.
- . (1986), 'The Philosophy of Logical Atomism', The Philosophy of Logical Atomism and Other Essays: 1914-19. Ed. John G. Slater. London: Allen. 160-244.
- Ryle, Gilbert (1949), The Concept of Mind. London: Hutchinson, 1966.
- Shoemaker, Sydney (1990), 'Qualities and Qualia: What's in the Mind?', Philosophy and Phenomenological Research 50, Supplement, Fall: 109-31.
- Sidgwick, Henry (1874), The Methods of Ethics. 5th ed. London: MacMillan, 1893.
- Sikora, Richard I. (1962), 'Reason, Language, and Choice: a critical analysis of contemporary development in ethical theory', Diss. Berkeley.
- . (1978), 'Is It Wrong to Prevent the Existence of Future Generations?', Obligations to Future Generations. Eds. R.I. Sikora and Brian Barry, Philadelphia: Temple UP. 112-66.
- . (1981), 'Synthetic A Priori Truths in an Artificial Language' Philosophy Research Archives 7.1449: 443-60.
- Simpson, George Gaylord (1969), Biology and Man. New York: Harcourt.
- Sleigh, Nicholas (1984), 'Brandt and Daniels on Credence in Moral Belief', Diss. Simon Fraser University, Essay I.
- . (1992), 'Objective Goodness and Aristotle's Dilemma', Journal of Value Inquiry 26: 341-51.
- . (1993), The Bion: the fundamental unit of evolution by natural selection. Unpublished ms., draft of May 1.
- Smart, J.J.C. (1956), 'Extreme and Restricted Utilitarianism', Philosophical Quarterly 6: 344-54.
- . (1973), 'An outline of a system of utilitarian ethics', Utilitarianism: for and against. By J.J.C. Smart and B. Williams. Cambridge: Cambridge UP. 3-74.

- Smart, R.N. (1964), Philosophers and Religious Truth. London: S.C.M. Press.
- Spinoza, Benedictus de (1967), Ethics. Trans. Andrew Boyle. Revised G.H.R. Parkinson. London: Dent.
- Strawson, Peter F. (1949), 'Ethical Intuitionism', Philosophy 24: 23-33.
- . (1982), 'Freedom and Resentment', Free Will. Ed. Gary Watson. New York: Oxford UP. 57-80.
- Szalay, Frederick S., and Eric Delson (1979), Evolutionary History of the Primates. New York: Academic.
- Taylor, Charles (1989), 'The Diversity of Goods', Anti-Theory in Ethics and Moral Conservatism. Eds. Stanley G. Clarke and Evan Simpson. Albany, New York: State U of New York P. 223-40.
- Timberlake, William, and Lucas, Gary A. (1989), 'Behavior Systems and Learning: From Misbehavior to General Principles', Contemporary Learning Theories: Instrumental Conditioning Theory and the Impact of Biological Constraints on Learning. Eds. Stephen B. Klein and Robert R. Mowrer. Hillsdale, New Jersey: Lawrence Erlbaum. 237-75.
- Tocqueville, Alexis de (1840), Democracy in America. Vol. 2. New York: Random House, 1990. 2 vols.
- Trivers, Robert (1985), Social Evolution. Menlo Park, California: Benjamin/Cummings.
- Williams, C.J.F. (1969), 'Are Primary Qualities Qualities?', Philosophical Quarterly 19: 310-23.