Metaethics

Judgments to the effect that certain things (or certain classes of things) are good or bad, right or wrong, or just or unjust, are first-order ethical judgments. Metaethics addresses second-order questions about the meaning and status of moral judgments, for example, “What does it mean to say that something is good or bad, or right or wrong?”, “Are moral judgments statements that purport to be true or false?”, and “In what sense, if any, can moral judgments be true or false (or correct or incorrect)?” Metaethical questions have been discussed throughout the history of philosophy, but systematic work on metaethics began early in the twentieth century with the publication of G. E. Moore’s *Principia Ethica*.

The first half of this entry discusses theories about the meaning of moral judgments, specifically, Moore’s theory, the Brentano-Ewing theory, emotivism, Hare’s prescriptivism, Foot’s theory, normative relativism, and Gibbard’s expressivism. The second half addresses the question of whether moral judgments are objectively true or false; it explains and assesses (some of) the main arguments for and against the view that moral judgments are objectively true or false. Questions about the truth of moral judgments are distinct from questions about moral knowledge. If moral judgments aren’t true or correct, then there is no such thing as moral knowledge. Moral knowledge is possible only on the assumption that there is something to know (i.e., moral truths). However, the view that there are objective moral truths does not imply that we have
knowledge of them; it is compatible with moral skepticism, the view that there are objective moral truths, but we cannot know what they are.

I. Theories of Meaning

Twentieth Century work on metaethics begins with questions about the meaning of moral judgments.

1. Moore’s Open Question Argument. In *Principia Ethica* Moore claims that the concepts of intrinsic goodness and badness are the most fundamental moral concepts. He says that the concepts of right and wrong can be defined in terms of "good" and "bad."

Moore writes:

> to assert that a certain line of conduct is, at a given time, absolutely right or obligatory, is obviously to assert that more good or less evil will exist in the world, if it be adopted than if anything else be done instead (Moore, p. 77).

Moore argues that goodness (the property denoted by the word “good”) is indefinable. By this he means that the property of goodness cannot be analyzed into constituent properties or elements. Goodness is a simple ultimate property like the property of being yellow, as opposed to a complex property such as the property of being a horse. Being a horse is analyzable in terms of other constitutive properties such as having a head, a heart, four legs, four hooves, and so on. Moore defends the claim that goodness is a simple unanalyzable property with his “open question argument.” This argument can be summarized as follows. Consider any definition of good or goodness according to which goodness is identical with a complex property (P). It will always make sense to ask if P is
good (it is an open question whether \( P \) is good, since it is not self-contradictory to deny that \( P \) is good). However, it makes no sense to ask whether \( P \) is \( P \) (this is not an “open question” - it is self-contradictory to deny that \( P \) is \( P \)). Therefore, goodness cannot be identical with \( P \). Moore considers several specific definitions of good which he subjects to the open question argument. He considers the view that good means pleasure, the view that good means what we desire, and the view that good means what we desire to desire. Clearly the question “Is pleasure good?” is not equivalent to the question “Is pleasure pleasant?” The statement “That which we desire to desire is good” is not equivalent to the statement “that which is good is good.” Moore’s open question test seems to work very nicely for these and many other definitions of good, but he gives no reason to think that every possible definition of goodness fails his open question test. Another serious problem with Moore’s argument is that he assumes that goodness cannot be identical with a property \( P \) unless the statement that \( P \) is good is analytic or true by definition. It is analytic (true by definition) that pleasure is pleasant, but it is not analytic that pleasure is good. Many contemporary philosophers contend that this assumption has been refuted by Hilary Putnam. According to Putnam, certain natural properties are identical with each other, even though statements to the effect that the properties are identical are not analytic (not true by definition). For example, water is H2O (the property of being water is identical with the property of being H2O), even though the statement that water is H2O is not true by definition.

2. Moore’s positive views. Moore is a “cognitivist” (i.e., he holds that moral
judgments are statements that ascribe properties to things). By contrast, noncognitivism is the view that moral judgments are not statements that ascribe properties to things. Moore claims that goodness is a simple unanalyzable property. Goodness is not a natural property like redness that can be perceived or apprehended through the five senses. Nonetheless, Moore claims that we can have direct intuitive knowledge of this property. This view is problematic. It is open to debate whether any such property exists. The quality of goodness that Moore posits is elusive; it’s difficult to know what he is referring to. Many people, on careful introspection, report that they do not intuit any such property. Another serious problem for Moore’s view is that it seems to be unable to account for the fact that moral judgments give or purport to give us reasons to act in certain ways as opposed to others. To say that something is intrinsically good implies that we have reasons to choose or prefer it. But it is unclear why this should be so if Moore’s theory is true. It is not clear why we should care whether or not our actions produce or fail to produce instances of the nonnatural properties that Moore describes and claims are identical to the properties of goodness and badness. (Even if they exist, it’s not clear that these nonnatural properties are “reason providing” in the way that something must be in order to be the property of goodness or badness.)

Moore's open question argument is one of the most influential arguments in the history of philosophy. Noncognitivist ethical theories, such as emotivism and prescriptivism, arose in a context of philosophical debate in which it was widely assumed that: (a) Moore has shown that goodness is indefinable, and (b) Moore's own positive
view is untenable - "good" does not refer to a simple nonnatural property that we directly intuit. Some philosophers concluded that moral terms do not refer to any properties at all and that moral judgments are not statements that ascribe properties to things.

3. Brentano and Ewing. Franz Brentano and A. C. Ewing agree with Moore that moral terms refer to “nonnatural” properties, but they give a very different description of the nature of those properties. They hold that the most fundamental moral properties are nonnatural relational properties of "fittingness" or "appropriateness" that hold between objects/properties and attitudes about them. Ewing holds that the relation of fittingness is unanalyzable and that our apprehension of it is self-evident. In The Definition of Good he says: “Certain characteristics are such that the fitting response to whatever possesses them is a pro-attitude, and that is all there is to it,” (p. 172). According to Brentano, to be good is to be an appropriate (fitting) object of love and to say that one thing is better than another is to say that it is correct to prefer it to the other. Ewing holds that to say that something is good means that it ought to be the object of a favorable attitude. This theory arguably avoids the second objection to Moore’s theory noted above. According to Brentano and Ewing, it is part of the meaning of moral judgments, (e.g., judgments to the effect that something is good or bad, or right or wrong) that they claim to give us reasons to have favorable or unfavorable attitudes about certain things and reasons to choose certain things over others.

4. Emotivism. Emotivism is the view that moral judgments are expressions of attitudes rather than statements that ascribe properties to things. Favorable (unfavorable)
moral judgments about something express favorable (unfavorable) attitudes about it. “Lincoln was a good man” means roughly "Yea Lincoln." So understood, emotivism denies the obvious phenomenon of moral disagreement. Suppose that you claim that Stalin was a good man and I claim that he was a bad man. If moral judgments were mere expressions of attitudes, then this could not constitute a disagreement. We might both agree that you like Stalin and I dislike him. Similarly, it is not a disagreement if you express your fondness for a particular flavor of ice cream and I express my distaste for that same flavor.

Ayer and Stevenson defend more sophisticated versions of emotivism. According to Ayer, to make a moral judgment is to express an attitude with the intention of influencing the attitudes or actions of other people - “Lincoln is a good man” means roughly “Yea Lincoln, catch the wave.” In cases of moral disagreement each party is attempting to alter the attitudes of the other. Stevenson holds that moral disagreement involves a disagreement in attitudes (the parties to the disagreement have incompatible attitudes about something) and each party is attempting to change the attitudes of the other party about the thing in question. Stevenson says that "x is good" means roughly "I approve of x; do so as well."

These revised versions of emotivism still do not afford a satisfactory account of moral disagreement. In cases of moral disagreement, people not only disagree in their attitudes and try to cause others to share their attitudes, they assert that their own attitudes are correct or justified and that the attitudes of those who disagree with them are
mistaken. If two people disagree about whether or not Stalin was a good man, each claims that the other’s attitudes about Stalin are mistaken or inappropriate.

Ayer on Moral Reasoning. Moral disputes often involve disagreements about factual questions. Ayer says that, to the extent that moral disagreements involve disagreements about “factual” questions, they can be rationally debated. For example, we can rationally debate whether the institution of capital punishment deters murder and whether it frequently results in the execution of innocent people. Sometimes, however, moral disagreements are based on differences in basic moral principles. Utilitarians believe that we should always do whatever will have the best consequences. Some people are unconditional pacifists. They believe that killing people is always wrong no matter what, even if killing saves many lives and produces much better consequences than not killing. Utilitarians and unconditional pacifists accept incompatible basic moral principles. According to Ayer, when people disagree about matters of basic principle, their disagreements cannot be rationally debated or rationally resolved.

[Allan Gibbards’s Wise Choices and Apt Feelings (discussed below) is a recent development of emotivism.]

5. R. M. Hare’s Prescriptivism. One of the notable features of this theory is that Hare offers a systematic reply to Ayer’s claims about the limits of moral reasoning. Hare claims that moral judgments are prescriptions that are universalizable and overriding. Prescriptions are commands, or imperatives, e.g., "Don't lie!" and “Shut the door!” Since commands aren’t statements that are true or false, prescriptivism is a non-cognitivist
theory. To say that moral judgments are universalizable means that, if one makes a moral judgment about a particular case, then one must make the same judgment about any cases that are similar to it in all morally relevant respects. (If I say that it is morally permissible for me to lie to my customers in a certain situation, then I am committed to the view that it would be permissible for others to lie to me in relevantly similar situations.) To say that moral judgments are overriding means that a person who makes moral judgments takes the prescriptions expressed by them to override any conflicting nonmoral considerations, such as considerations of prudence, etiquette, and the law. According to Hare, statements of the form "It is morally wrong all-things-considered for you to do X, but, nevertheless, you would be justified in doing X," are self-contradictory. On Hare’s view, it is also inconsistent to say that it is morally wrong (all things considered) for you to do X but still command or advise you to do X.

In response to Ayer’s claims about the limits of moral reasoning, Hare would say that sometimes we can argue against another person’s moral judgments by showing that they are inconsistent. Hare claims that requirements of consistency severely constrain the kinds of moral judgments we can make. Suppose that a dishonest plumber claims that it is morally permissible (or even obligatory) for him to defraud his customers and bill them for unneeded repairs that cost them thousands of dollars. To be consistent, he must say that it would be morally permissible (obligatory) for others to defraud him and those he cares about in relevantly similar (hypothetical and/or actual) cases. Since moral judgments are prescriptive, he is committed to prescribing that others defraud him and
those he loves in relevantly similar circumstances. Consistency also requires that he refrain from objecting if others defraud him and those he loves in such cases.

In *Freedom and Reason*, Hare considers the case of a Nazi who claims that it is his moral duty to kill Jews. To be consistent, the Nazi must hold that others should kill him if he is Jewish. Suppose that we show the Nazi that, unbeknownst to him, he and his wife are Jewish. To be consistent, the Nazi must say (and mean) “All right send me and my family to an extermination camp.” Since moral judgments are overriding, the Nazi can’t consistently make any commands or pleas to the contrary. Hare thinks that few people can be consistent Nazis. Hare allows that a Nazi could be consistent if s/he so hates Jews that s/he sincerely holds that [in Hare’s words] “‘Jews are such an abomination that I and my whole family, if we were Jews, should be sent to the gas chamber’” (Hare 1963, p. 172). One can be a consistent Nazi if one is willing to have one’s moral principles applied against one’s own interests and the interests of those one loves. Hare gives no reason to think that the Nazi’s distinctive moral views are false or mistaken, just that it is difficult to be a consistent Nazi. This concession bodes ill for Hare’s theory of moral reasoning, since if it can’t establish the correctness of the view that the Nazi's actions are wrong, it is doubtful that it can establish the correctness of any ethical judgments.

However, Hare is too quick to concede the limits of his arguments in this case. It may be possible to be a consistent Nazi, provided that one has a very great hatred of Jews and desires their extermination more than the continuation of one’s own life and the lives of one’s loved ones. The obvious question to ask here is whether such hatred is rational, and
it seems that it is not. Such hatred depends on numerous false beliefs about the characteristics of Jews and their responsibility for the ills of the world. A Nazi could be consistent provided that she is willing to have her principles applied against herself and her loved ones, but a Nazi couldn’t be both consistent and adequately informed about matters relevant to her moral convictions. (See Hare’s Moral Thinking for a later development of his views.)

As Hare himself notes, his consistency arguments apply only to people who make moral judgments. They don’t apply to amoralists who refrain from making moral judgments. Hare’s consistency arguments can’t show why we shouldn’t be amoral. Even more worrisome for Hare’s purposes is that people who employ alternative normative concepts are able to endorse horrendous acts such as the extermination of the Jewish people. Suppose that a Nazi rejects the concepts of morally right and wrong actions in favor of a “Code of Honor” according to which it is “honorable” to kill Jews. Hare’s theory doesn’t give us any basis for criticizing such views.

This last possibility raises an important and somewhat neglected set of issues. There are many different alternative normative concepts (e.g., concepts of moral obligation and right and wrong, concepts of virtue, and concepts of honor). People are free to employ any of these concepts and order their lives in accordance with them; people are also free to reject any of these concepts. Philosophers who write about metaethics need to say much more about the choices we make in accepting and rejecting various normative concepts. They also need to say much more about the question of how,
if it all, we can justify the choices we make in accepting/wielding certain concepts rather than others. [See Nietzsche, Hare, Blackburn, Mackie, and Williams for discussions that shed light on these issues; also see the discussions of Foot, Gibbard, and incommensurability below.]

6. Foot and Thick Moral Concepts. Both emotivism and prescriptivism imply that the concepts of good and bad, and right and wrong have no fixed descriptive meaning. One can consistently apply these terms to any things (or any actions). For example, it is perfectly consistent to say that it is morally obligatory to clasp and unclasp one’s hands every half hour or to say that bringing it about that the number of hairs on one’s head is an even number is a great intrinsic good. These are sincere coherent ethical judgments provided that the person who makes them has the attitudes they express or is willing to consistently universalize the prescriptions that they express. Philippa Foot argues that this is a serious mistake, because there are limits to the things to which "good" and other moral terms can be consistently applied. She thinks that emotivism and prescriptivism make a mistake that is comparable to the sort of mistake one would make if one said that being proud of something consists simply in having a certain sort of attitude about it and that, in principle, anything could be the object of one’s pride. Pride is not just a feeling, a welling up in one's chest that one can have about anything. The object of one’s pride must be (a) one's own somehow and (b) an achievement or something that one takes to be good. Even if I puff up my chest and feel it welling up as I look at the sky, it is not correct to say that I am (feeling) proud of the sky unless there is some background belief
that explains why I think that it is somehow mine. Foot claims that something can be called a "good action" only if it satisfies one of the following conditions: (a) it is the fulfillment of a special duty derived from a role or promise, or (b) it exemplifies a virtue. It follows that we can't say that twiddling one's thumbs four times each day, for instance, is a morally good action in the absence of special reasons for thinking that it fulfills a duty or exemplifies a virtue.

Foot’s own analysis of the meaning of moral judgments is a combination of noncognitivism and (naturalistic) cognitivism. She claims that moral judgments have both evaluative meaning (they express attitudes and guide actions) and descriptive meaning. Moral concepts which have both kinds of meaning are called “thick concepts.” emotivism and prescriptivism claim that the concepts of right and wrong, and good and bad are “thin” concepts. Thin normative concepts have no fixed descriptive meaning, only evaluative meaning.

Foot’s theory aptly describes the meaning of the terms we use to refer to moral virtues and vices. Terms such as "generous," "cowardly," and "honest" are "thick" concepts - they have both evaluative and descriptive meaning. The words "generous" and "honest" commend or express favorable attitudes about the things to which they are applied. There are clear descriptive criteria for using such terms. It is a misuse of language to apply them to things that do not satisfy those criteria. It would be a misuse of the word "generous" to apply it to someone who never gives any tangible goods or time or effort to other people, even though s/he has a great deal of money and leisure time and
many opportunities to help others in need. Foot’s theory helps us to frame some important questions - “In ordinary language, are the terms ‘good’ and ‘bad’ and ‘right’ and ‘wrong’ thick or thin concepts?” “If these concepts are thin concepts, should we dispense with them in favor of thick concepts?” Foot seems to think that the concepts of good and bad, and right and wrong are thick concepts. By contrast, some proponents of “virtue ethics,” including Bernard Williams, think that the concepts of good and bad and right and wrong are thin concepts. However, they think that these thin concepts should be dropped or greatly downplayed in favor of the thick concepts that refer to virtues and vices.

Thick terms mandate particular evaluations of certain kinds of things. Many thick concepts encapsulate objectionable evaluations (e.g., ethnic slurs). The word "n____" only applies to people of African origin; it cannot be correctly applied to Chinese or Europeans. The word "n____" also expresses contempt for Africans. Those who don't think that Africans, qua Africans, are worthy of contempt do not use the word "n____" (or don’t use the word non-ironically). (Similar comments apply to all other “ethnic slur terms.”) Honor is another example of a thick concept that many people have reasonably chosen to abandon because they reject the evaluations implicit in its use. Given certain concepts of honor, it is dishonorable for me to not to challenge you to a duel to the death if you insult me or show me disrespect. Given other concepts of honor, it is dishonorable for me not to kill my sister if she is raped. These examples make it clear that, for any thick concept that we employ, we should be open to criticisms of the evaluations implicit
in that concept and consider the possibility that they are mistaken and cease employing the concept. Thus, it is at least arguable that we need higher-level “thin” concepts in terms of which to assess the evaluations implicit in the thick concepts we use and encounter.

7. **Normative Relativism.** This theory is defended by many anthropologists, including, Ruth Benedict and William Sumner (in Moser and Carson, pp. 69-89). They claim that “X is morally right” means roughly “X is approved of by my society.” This view is open to very serious objections. It implies that statements such as "Slavery is morally wrong, even though my society approves of slavery" are self-contradictory. But such statements are not self-contradictory. A person can criticize or dissent from the moral standards of her own society without contradicting herself. Normative relativism also implies that many ostensible moral disagreements between members of different societies are not genuine disagreements. Suppose that I am a member of a society that approves of the institution of slavery and you are a member of a society that disapproves of slavery. I claim that slavery is just and morally permissible. You object and claim that slavery is unjust and impermissible. Surely this is a moral disagreement; what I say contradicts what you say. But, according to normative relativism, there is no disagreement in this case. My statement is perfectly consistent with your statement and both statements are true - it is true that my society approves of slavery and true that your society disapproves of slavery.

8. **Gibbard’s Expressivism.** This theory is a recent development of emotivism that
incorporates elements of the Brentano-Ewing theory. Gibbard analyzes moral judgments as claims about the rationality (or aptness) of feelings of guilt and anger.

What a person does is **morally wrong** if and only if, it is rational for him to feel guilty for having done it, and for others to be angry at him for having done it (Gibbard, p. 42)

He defends an expressivist/emotivist analysis of rationality. According to Gibbard, to say that something (an act, belief, or feeling) is rational is to express one's acceptance of norms that permit it. Unlike Moore, Hare, and Foot, Gibbard does not claim to be offering an analysis of the ("ordinary language") meaning of moral terms. Rather, he describes his theory as a proposal about how to use normative concepts. Evidently, not every society has a normative system that includes norms for guilt and anger. Thus, on Gibbard’s narrow construal of "morality," not every society has a moral code. Gibbard raises important questions about our choices between alternative normative concepts. Among other things, he asks about the value of morality (narrowly construed) - would we be better off with a normative code in which norms for guilt and anger didn’t play a central role? Gibbard offers an answer to Nietzschean criticisms about the value of morality. He says that moral norms help coordinate guilt and anger. Guilt assuages anger and thereby helps promote peace between human beings. Normative codes that don't include norms for guilt, will not be able to assuage anger and promote reconciliation between human beings as well as moral codes.

**II. Moral Truth, Moral Reality**
We now turn to questions about moral truth and moral reality. In what sense, if any, can moral judgments be true or false, or correct or incorrect? Are moral judgments objectively true or false in the way that we take ordinary “factual” statements to be? Is there a moral reality or something else in virtue of which moral judgments are true or false (or correct or incorrect)? We can’t answer these questions simply by appealing to theories of meaning. When we ask whether moral judgments are true or false we are not simply asking about what we mean or claim when we make moral judgments. We are asking whether there is anything that backs up our moral judgments and makes them true or correct.

Cognitivism, Non-Cognitivism, and the Truth of Moral Judgments. Cognitivists hold that moral judgments are statements that purport to be true. This is compatible with the view that there are no moral facts (no moral reality) that back up our moral judgments and make (some of them) true or correct. J. L. Mackie holds such a view, which he calls an “error theory” of morality. Mackie is a cognitivist who claims that moral judgments are statements that assert or presuppose the existence of objective values. However, he claims that, since objective values don’t exist, all moral judgments are false.

Non-cognitivists hold that moral judgments are not statements that purport to be true or false. Strictly speaking, non-cognitivists can’t say that moral judgments are true or false. However, they can still say that moral judgments possess something that closely resembles truth or falsity. Emotivists can say that moral judgments are reasonable or unreasonable depending on whether the emotions or attitudes they express are reasonable
or unreasonable. (At the very least, attitudes and emotions can be unreasonable if they are based on false beliefs.) Prescriptivists can also make sense of something resembling the idea of moral truth. In Moral Thinking, Hare claims that there are certain moral judgments that an informed consistent person must endorse, provided that s/he makes any moral judgments at all. These are judgments that we can reject only by opting out of moral discourse altogether.

**Moral Objectivism.** Our ordinary notion of truth is a notion of **objective truth.** If something is true, then it is **true for everyone** (and true for everyone, everywhere, at all times). [Thus, it is misleading to use the word “truth” as many relativists do when they claim that the truth of moral judgments is “relative to” different people so that a moral judgment that is “true for” one person may not be “true for” another.] Let us use the term “objectivism” to refer to the view that moral judgments are objectively true or false (or objectively correct or incorrect in some sense that closely resembles truth or falsity). We should distinguish between the view that there is an objectively correct answer to every moral question and the view that there are objectively correct answers to some, but not all, moral questions. Call the former view “unqualified objectivism” and the latter view “qualified objectivism.” Call the view that there are no objectively correct answers to any moral questions “unqualified non-objectivism.” Unqualified objectivism implies that for any moral question, e.g., “Was it right for Ms. Jones to have an abortion in April 1999?,” there is an objectively correct answer and that anyone who gives a conflicting answer is mistaken. Qualified objectivism holds that there are some moral questions...
about which there are objectively correct answers and other moral questions about which there are no objectively correct answers. Unqualified nonobjectivism implies that for, any moral question, there is no objectively correct answer to that question.

**Moral Realism.** Statements such as “the earth is less than 100,000,000 miles from the sun,” are true in virtue of facts that hold independently of what we believe or desire. Moral realism is the view that there are moral facts in virtue of which moral judgments are objectively true or false and that these facts are logically independent of the beliefs, attitudes, emotions, or preferences of rational beings and independent of the beliefs, attitudes, emotions, or preferences that rational beings would have in hypothetical situations (e.g., the moral beliefs that someone would have if s/he were fully informed about relevant facts). Moral nonrealism is the view that there are no independent moral facts. The truth of moral realism would guarantee the truth of moral objectivism, but one can be a moral objectivist without being a moral realist. Kant, Roderick Firth, and Michael Smith (see below) are moral objectivists but not moral realists. Hare’s *Moral Thinking* also defends non-realist moral objectivism.

### IIa. Arguments Against Moral Objectivism

We now turn to arguments against moral objectivism.

1. **Disagreement** Moral disagreement is widespread among ostensibly sane and rational people. Consider the following argument:

   1. There is disagreement among rational people about the answers to all (some) moral questions.
2. If there is disagreement among rational people about the answer to a question, then there is no objectively correct or objectively true answer to that question.

Therefore,

Unqualified (qualified) nonobjectivism is true. (There are no moral questions (there are only some moral questions) for which there is an objectively correct answer.)

The cogency of this argument depends on the account of moral truth or correctness that the objectivist gives. If moral realism is true, then there are moral facts in virtue of which moral judgments are objectively true that are independent of what we believe. So, if moral realism is true, then moral objectivism is true and the phenomenon of moral disagreement among rational people isn’t a serious objection to moral objectivism. Similarly, disagreement between reasonable people does not constitute any kind of objection to the view that ordinary historical judgments are objectively true or false.

Consider the question “Did Lee Oswald fire any of the shots that killed President Kennedy?” Rational people disagree about the answer to this question, and, at the present time, it may be impossible to know for certain what the answer is. In spite of the disagreement, there is an objective fact of the matter - either Oswald fired some of the fatal shots or else he did not. However, the phenomenon of ethical disagreement among ostensibly rational and well-informed people constitutes a serious argument against attempts to defend moral objectivism by appeal to theories of rationality, because such theories claim that objective moral truths are constituted by the agreement of rational people. One standard rejoinder to the objection about disagreement is the claim that
“ideally rational” or “fully rational” people would not have moral disagreements. (See the discussion of the ideal observer theory below.)

**Digression: Disagreement vs. Incommensurability.** Moral disagreement should not be confused with moral incommensurability. Two people can disagree about whether an act is right or wrong only if they share the concepts of right and wrong action. Often, the differences between the moral views of different societies constitute cases of incommensurability rather than disagreement. Many philosophers and cultural anthropologists (including Nietzsche and Anscombe) claim that the concept of moral obligation is unique to the Judaism, Christianity, Islam, and the civilizations that developed from those religions. If this is true, then it is doubtful that Ghengis Khan and his warriors possessed our concept of a moral obligation and morally right and wrong actions. I condemn the Mongol destruction of Iraq in the 13th Century as a morally wrong action. Many or most of the Mongols who took part in this did so in good conscience, but it would probably be incorrect to say that they thought that what they did was morally right (they couldn’t have this belief unless they employed the concepts of right and wrong actions). Even though we don’t disagree about the moral rightness of this action, there is surely some kind of disagreement here. My moral judgments and the attitudes they endorse are contrary to many of Ghengis Khan’s attitudes; Ghengis Khan clearly approves of (and thinks it correct to approve of) actions that I disapprove of and condemn as morally wrong.

2. **The Appeal to Tolerance.** This argument goes roughly as follows. We should
be respectful and tolerant of other people’s moral views, however, moral objectivism implies that many people's views are mistaken and thus not worthy of respect or toleration. This argument is widely accepted and motivates many people to reject moral objectivism, but, on examination, it is a very weak argument that few philosophers take seriously. First, endorsing moral objectivism does not commit one to being intolerant of the moral views of others. If I am an unqualified objectivist, then I think that there are objectively correct answers to every moral question. However, my being an objectivist does not entail that I claim to know what those answers are. Nor does my being an objectivist entail that I think that the views of others who disagree with me are worthy of disrespect or suppression. An objectivist can claim that objectively true moral principles require tolerance and respect for the views of others. Second, non-objectivism doesn't imply that we should be tolerant. Non-objectivists can endorse first-order moral principles that permit or require them to be intolerant of the views of others. All that follows from nonobjectivism is that one's moral judgments, whatever they happen to be, aren't objectively true or false or objectively correct or incorrect.

3. **Moral Explanations and Moral Facts.** Gilbert Harman argues that it is unnecessary to posit the existence of moral facts in order to explain phenomena. Thus, moral facts are superfluous entities - there is no reason to suppose that they exist.

   [O]bservation plays a role in science that it does not play in ethics. The difference is that you need to make assumptions about certain physical facts to explain the occurrence of the observations that support a scientific
theory, but you do not seem to need to make assumptions about any moral facts to explain the occurrence of the so-called moral observations I have been talking about. In the moral case, it would seem that you need only make assumptions about the psychology or moral sensibility of the person making the moral observation. In the scientific case, theory is tested against the world (Harman, p. 6).

Among the phenomena that moral facts might explain are the moral judgments we make and the moral sentiments we feel (e.g., feelings of guilt and indignation). According to Harman, these phenomena are fully explained by our psychology and the fact that we accept certain moral principles; we don't need to assume the existence of moral facts or assume that those principles are true. Harman gives the following example. Someone tortures an animal. You believe that this action is wrong and you feel moral indignation. Harman says that we don't need to postulate moral facts in order to explain your belief and your indignation. They are explained by the fact that you were taught certain moral principles and have a certain psychological make up. Your accepting the moral principles in question is necessary to explain your beliefs and your indignation, but we don’t need to assume that your moral principles are true. By contrast, in science, we can justify the postulation of entities by their ability to explain our observations of the world. The postulation of atoms helps to explain such things as Geiger counters and nuclear bombs. (Nicholas Sturgeon offers an influential reply to this argument - see below.)
IIb. Arguments in Favor of Moral Objectivism

Most contemporary philosophers who defend moral objectivism do so on one of the following three grounds.

1. The Appeal to Moral Realism. As explained earlier, the truth of moral realism would guarantee the truth of moral objectivism. Some versions of moral realism claim that moral properties are “nonnatural” properties. Such theories are widely criticized on the grounds that the entities that they postulate don’t exist or that those entities can’t plausibly be identified with moral properties (see above). In light of these criticisms of nonnaturalist versions of moral realism, recent attempts to defend naturalistic versions of realism are particularly noteworthy.

Sturgeon’s Naturalistic Realism. Sturgeon claims that moral facts are constituted by natural facts. Sturgeon holds that moral properties are identical with natural properties, but he does not take statements asserting the identity of moral properties and natural properties to be analytic. He claims that his view is invulnerable to Moore’s open question argument. Sturgeon also attempts to answer Harman’s argument about explanation. He defends his theory on the grounds that moral facts help explain certain phenomena. Sturgeon offers the following example. Hitler’s moral depravity helps explain why he started WWII and ordered the Holocaust. If Hitler hadn't been morally depraved, he wouldn't have started WW II and he would not have ordered the Holocaust.

Harman's thesis implies that the supposed moral fact of Hitler's being morally depraved is irrelevant to the explanation of Hitler's doing what he
did.... To assess this claim, we need to conceive a situation in which Hitler was not morally depraved and consider the question whether in that situation he would still have done what he did. My answer is that he would not, and this answer relies on a (not very controversial) moral view: that in any world at all like the actual one, only a morally depraved person could have initiated a world war, ordered "the final solution," and done any number of other things Hitler did. That is why I believe that if Hitler hadn't been morally depraved, he wouldn't have done those things, and hence that the fact of his moral depravity is relevant to an explanation of what he did (Sturgeon, “Moral Explanations,” in Sayre-McCord, p. 249).

Sturgeons’s arguments have generated a very lively debate. (Criticisms by Terrance Horgan and Mark Timmons are particularly noteworthy.)

2. Theories of Rationality. Many nonrealists claim that there are objective moral facts that are constituted by facts about what it is rational for people to believe or desire or “will.” Kant, Hare, and Korsgaard defend such views. Kant holds that moral truths (truths about what is right and wrong) are truths about what we can rationally and consistently will. For Kant moral truths are truths of reason.

The ideal observer theory (IOT) uses the idea of an ideally rational moral judge or ideal observer as a standard for the truth of moral judgments. According to Roderick Firth’s version of the IOT, a favorable moral judgment about X ("X is good/right") is (objectively) true provided that all ideal observers would feel approval for X. An
unfavorable moral judgment about X ("X is bad/wrong") is (objectively) true provided that all ideal observers would feel disapproval for X (Firth, “Ethical Absolutism and the Ideal Observer,” in Hospers and Sellars, pp. 200-221). In Ethical Theory, Richard Brandt defends a different version of the IOT. Brandt says that a moral judgment X is objectively true provided that all ideal observers would accept or believe X. (David Hume and Adam Smith also defend versions of the ideal observer theory.) Firth ascribes the following characteristics to an ideal observer: (a) omniscience or knowledge of all nonmoral facts; (b) omnipercipience, or the ability to imagine vividly any events or states of affairs, including the experiences of others; (c) disinterestedness, i.e., not having any interests or desires that involve essential reference to particular persons or things; (d) dispassionateness, i.e., not having any emotions that are directed upon objects because they are believed to have essentially particular features; (e) consistency; and (f) normality "in other respects."

Firth thinks that all ideal observers would feel approval and disapproval for the same things. Given this, and given his version of the ideal-observer theory, unqualified moral objectivism is true. Brandt thinks that ideal observers would agree about the answers to some, but not all, moral questions. He thinks that the IOT commits us to qualified objectivism. If both Firth and Brandt are mistaken and ideally rational moral judges could disagree in their attitudes or judgments about every moral question, then the IOT commits us to unqualified nonobjectivism.

Ideal observers are characterized as “informed” or “fully informed.” Brandt says
that ideal observers must possess all information “relevant to” the issues they judge. Firth notes difficulties in determining which facts are and are not relevant to answering a given moral question. He contends that there is no way to determine which facts are and are not relevant to a given moral question without presupposing answers to controversial moral questions that the IOT is supposed to provide rather than presuppose. Because of this, he feels compelled to say that an ideal observer is omniscient with respect to all non-moral facts. There is an unintended irony in Firth’s characterization of an ideal observer as a human being who omniscient, "but otherwise normal." Humans are very far from being omniscient. It is not clear that it makes sense to talk about how you or I would react if we were omniscient. An omniscient being would have to be God or some kind of deity. If we press this point, then the IOT starts to look a lot like a divine will theory of morality.

Michael Smith’s The Moral Problem defends a theory that closely resembles the IOT. With qualifications, Smith holds that to say that an action is morally right means that we have normative reason to do it. What a person has normative reason to do is what s/he would desire to do if s/he were fully rational (being fully rational includes having no false beliefs and having all relevant true beliefs). Smith stresses that what I have normative reason to do (in my actual circumstances) depends on what my fully rational self would want (or advise) my actual self to do. (Smith’s theory is sometimes called an “ideal advisor theory.”) Smith thinks that his theory implies that there are (objective) moral facts. On his account, the judgment “it is right for S to do X” is objectively true provided that we would all desire that S do X if we were fully rational. Smith thinks that
an ideally rational process of reflection and debate between people who initially disagree in their desires about what people should do is likely to yield (complete) agreement. According to Smith, thick moral concepts such as “honesty” and “treachery” reveal considerable agreement about what is right and wrong. The common use of such concepts reveals that nearly everyone agrees that acts of treachery are wrong (other things equal) and that acts of honesty are right (other things equal). Smith also argues that much seemingly intractable moral disagreement has its origins in (unreasonable) appeals to religious authority. (Others take a very different view about the relevance of religion to these issues; see below.) Smith says that the case for moral objectivism ultimately depends on the outcome of (first-order) debates in normative ethics. In order determine “whether or not there really are moral facts” we must “engage in normative ethical debate and ... see where the arguments that we give ultimately lead us” (Smith, p. 202). “The real question is whether we will, by engaging in such debate, come up with answers to moral questions that secure the free agreement of those who participate” (Smith, p. 201).

3. The Appeal to God's Will. Some hold that God’s will constitutes the objective standard for the truth of moral judgments. The view that God created human beings for certain purposes is one way of making sense of the widely held view that ethical theories should be based on theories of human nature or the “telos” of human life. The most well known theory that attempts to base (objective) morality on God’s will is the divine command theory. The traditional divine command theory (TDCT) holds that God's commands constitute the ultimate standard of right and wrong. An action is morally
obligatory if, and only if, God commands it. An action is morally permissible if, and only if, God permits it. An action is morally wrong if, and only if, God forbids it. What makes an act morally obligatory is that God commands it; what makes an act morally permissible is that God permits it; what makes an act is morally wrong is that God forbids it.

There are a several standard objections to the TDCT. These objections are widely regarded as fatal or decisive. i) The TDCT implies that nothing anyone does can be morally obligatory or morally wrong, unless God exists and commands and forbids us to do certain things. However, certain actions would be right or wrong even if God didn’t exist. ii) The TDCT implies that any act would be right if God commanded us to do it. But certain acts (e.g., acts of cruelty or murder) would be wrong even if God commanded us to perform them. iii) The TDCT implies that what is wrong/obligatory is wrong/obligatory because God forbids/commands it. The TDCT does not allow us to say that God forbids murder because murder is wrong. Rather, the TDCT implies that murder is wrong because God forbids it. Thus, the TDCT implies that God has no reason to command one thing rather than another and God’s arbitrary commands can’t be the basis for genuine moral obligations.

These may be fatal objections to the TDCT. However, it doesn’t follow that all theories that attempt to make God’s will the basis for an objective morality are subject to fatal objections. Robert Adams has formulated a modified version of the DCT that avoids all of the objections to the TDCT. Adams’s modified DCT can be stated roughly as
follows:

If there is a loving God then: 1) an action is obligatory if, and only if, a loving God commands it; 2) an action is morally permissible if, and only if, a loving God permits it; and 3) an action is morally wrong if, and only if, a loving God forbids it. If there does not exist a loving God, then the rightness or wrongness of actions is determined in some other way.

Adams holds that if there is a loving God, then right and wrong are determined by God's commands; if there does not exist a loving God, then right and wrong are determined in some other way. Thus, Adams's theory avoids the first objection. Adams's modified DCT also avoids the second objection. It does not imply that we would be morally obligated to obey God's commands if God commanded cruelty. If God commanded cruelty for its own sake he would thereby show himself to be unloving. Adams's theory does not imply that we would be obligated to follow the commands of a cruel or unloving God. Adams can't say that God commands what he commands because it is morally right (independently of being commanded by God). But Adams is not committed to the view that God’s commands are arbitrary or that God has no reason to command one thing rather than another. Adams can say that God commands what he commands because of his loving nature and because he is omniscient. (See Linda Zagzebski’s Divine Motivation Theory for a very different sort of religiously based moral theory.)

Suppose that there exists an omniscient loving God who created human beings for certain purposes. Suppose also that moral realism is false and there are no independent moral facts to which God’s will must conform on pain of error. Given all of this, it is plausible to regard God’s will and purposes as objective standards of morality; God’s
standpoint for assessing things is arguably more authoritative than that of a maximally rational human being. If moral realism is false and God doesn’t exist, then the most promising basis for defending moral objectivism is by appeal to a theory of rationality.

**Bibliography**


Anthologies (All of These Include Extensive Bibliographies)


