

PDF version of the entry
The Definition of Morality
<https://plato.stanford.edu/archives/fall2020/entries/morality-definition/>
from the FALL 2020 EDITION of the

STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Edward N. Zalta Uri Nodelman Colin Allen R. Lanier Anderson
Principal Editor Senior Editor Associate Editor Faculty Sponsor

Editorial Board
<https://plato.stanford.edu/board.html>

Library of Congress Catalog Data
ISSN: 1095-5054

Notice: This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

Stanford Encyclopedia of Philosophy
Copyright © 2020 by the publisher
The Metaphysics Research Lab
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305

The Definition of Morality
Copyright © 2020 by the authors
Bernard Gert and Joshua Gert

All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

The Definition of Morality

First published Wed Apr 17, 2002; substantive revision Tue Sep 8, 2020

The topic of this entry is not—at least directly—moral *theory*; rather, it is the definition of morality. Moral theories are large and complex things; definitions are not. The question of the definition of morality is the question of identifying the *target* of moral theorizing. Identifying this target enables us to see different moral theories as attempting to capture the very same thing. And it enables psychologists, anthropologists, evolutionary biologists, and other more empirically-oriented theorists to design their experiments or formulate their hypotheses without prejudicing matters too much in terms of the specific content a code, judgment, or norm must have in order to count as distinctively moral.

There does not seem to be much reason to think that a single definition of morality will be applicable to all moral discussions. One reason for this is that “morality” seems to be used in two distinct broad senses: a descriptive sense and a normative sense. More particularly, the term “morality” can be used either

1. descriptively to refer to certain codes of conduct put forward by a society or a group (such as a religion), or accepted by an individual for her own behavior, or
2. normatively to refer to a code of conduct that, given specified conditions, would be put forward by all rational people.

Which of these two senses of “morality” a moral philosopher is using plays a crucial, although sometimes unacknowledged, role in the development of an ethical theory. If one uses “morality” in its descriptive sense, and therefore uses it to refer to codes of conduct actually put forward by distinct groups or societies, one will almost certainly deny that

there is a universal morality that applies to all human beings. The descriptive use of “morality” is the one used by anthropologists when they report on the morality of the societies that they study. Recently, some comparative and evolutionary psychologists (Haidt 2006; Hauser 2006; De Waal 1996) have taken morality, or a close anticipation of it, to be present among groups of non-human animals: primarily, but not exclusively, other primates.

Accepting that there are two uses or senses of “morality”—a descriptive sense and a normative sense—does not commit one to holding that the “distinction between descriptions and norms—between what is and what ought to be—is obvious and unbridgeable”, as some have held that it does (Churchland 2011: 185). To see this, note that it is obvious that there is a descriptive sense of morality. That is, it is obvious that one can sensibly describe the moralities of various groups without making any normative claims. And it should be equally obvious that that one might hold that a certain code of conduct would be put forward by all rational people under certain conditions without having any particular views about the nature of the is/ought gap or the possibility of crossing it.

Any definition of “morality” in the descriptive sense will need to specify *which* of the codes put forward by a society or group count as moral. Even in small homogeneous societies that have no written language, distinctions are sometimes made between morality, etiquette, law, and religion. And in larger and more complex societies these distinctions are often sharply marked. So “morality” cannot be taken to refer to every code of conduct put forward by a society.

In the normative sense, “morality” refers to a code of conduct that would be accepted by anyone who meets certain intellectual and volitional conditions, almost always including the condition of being rational. That a person meets these conditions is typically expressed by saying that the

person counts as a *moral agent*. However, merely showing that a certain code would be accepted by any moral agent is not enough to show that the code is the moral code. It might well be that all moral agents would also accept a code of prudence or rationality, but this would not by itself show that prudence was part of morality. So something else must be added; for example, that the code can be understood to involve a certain kind of impartiality, or that it can be understood as having the function of making it possible for people to live together in groups.

As we’ve just seen, not all *codes that are put forward by societies or groups* are moral codes in the descriptive sense of morality, and not all *codes that would be accepted by all moral agents* are moral codes in the normative sense of morality. So any definition of morality—in either sense—will require further criteria. Still, each of these two very brief descriptions of codes might be regarded as offering some features of morality that would be included in any adequate definition. In that way they might be taken to be offering some *definitional* features of morality, in each of its two senses. When one has specified enough definitional features to allow one to classify all the relevant moral theories as theories of a common subject, one might then be taken to have given a definition of morality. This is the sense of “definition” at work in this entry.

Explicit attempts, by philosophers, to define morality are hard to find, at least since the beginning of the twentieth century. One possible explanation for this is the combined effect of early positivistic worries about the metaphysical status of normative properties, followed (or augmented) by Wittgensteinian worries about definitions of any significant terms whatsoever. Whatever the explanation, when definitions have been offered, they have tended to be directed at the notion of moral judgment (Hare 1952, 1981) rather than at morality itself. However, to the degree that these definitions of moral judgment are adequate, they might, without much effort, be converted into definitions of morality in the descriptive

sense. For example, a particular person's morality might be regarded as the content of the basic moral judgments that person is prepared to accept.

One might use a detailed definition of moral judgment to define morality in a descriptive sense in another way—other than simply as the content of a person's moral judgments, or the content of the moral judgments that prevail in a certain society or group. In particular, the very features of a judgment that make it qualify as a moral judgment might be transposed from a psychological key to something more abstract. Here is one simplified example. Suppose that a negative judgment of an action only counts as a negative moral judgment if it involves the idea that there is a *prima facie* case for punishing that action. In that case, a definition of morality in the descriptive sense will include a corresponding idea: that the prohibitions of morality, taken in the descriptive sense, are those that are backed by the threat of punishment. Of course, if one goes this route, other conditions will need to be included, to differentiate morality from criminal law.

What counts as definitional of morality, in either sense of "morality", is controversial. Moreover, the line between what is part of a definition, in the sense at issue, and what is part of a moral theory, is not entirely sharp. For example, some might regard it as definitional of morality, in the normative sense, that it governs only interpersonal interactions. Others, however, might take this to be a substantive theoretical claim. Some might take it as definitional of "morality" in its descriptive sense that it be a code of conduct that a person or group takes to be most important. But others might say that attention to religion casts doubt on this idea.

"Morality", when used in a descriptive sense, has an important feature that "morality" in the normative sense does not have: a feature that stems from its relational nature. This feature is the following: that if one is not a member of the relevant society or group, or is not the relevant individual,

then accepting a certain account of the content of a morality, in the descriptive sense, has no implications for how one thinks one should behave. On the other hand, if one accepts a moral theory's account of moral agents, and of the conditions under which all moral agents would endorse a code of conduct as a moral code, then one accepts that moral theory's normative definition of "morality". Accepting an account of "morality" in the normative sense commits one to regarding some behavior as immoral, perhaps even behavior that one is tempted to perform. Because accepting an account of "morality" in the normative sense involves this commitment, it is not surprising that philosophers seriously disagree about which account to accept.

- 1. Is Morality Unified Enough to Define?
- 2. Descriptive Definitions of "Morality"
- 3. Implicit and Explicit Definitions in Allied Fields
- 4. Normative Definitions of "Morality"
- 5. Variations
 - 5.1 Morality as linked to norms for responses to behavior
 - 5.2 Morality as linked to advocacy of a code
 - 5.3 Morality as linked to acceptance of a code
 - 5.4 Morality as linked to justification to others
- Bibliography
- Academic Tools
- Other Internet Resources
- Related Entries

1. Is Morality Unified Enough to Define?

An assumption suggested by the very existence of this encyclopedia entry is that there is some unifying set of features in virtue of which all moral systems count as moral systems. But Sinnott-Armstrong (2016) directly

argues against an analogous hypothesis in connection with moral judgments, and also seems to take this view to suggest that morality itself is not a unified domain. He points out that moral judgments cannot be unified by any appeal to the notion of harm to others, since there are such things as moral ideals, and there are harmless behaviors that a significant number of people regard as morally wrong: Sinnott-Armstrong gives example such as cannibalism and flag-burning. Whether people who condemn such behaviors morally are correct in those judgments is largely irrelevant to the question of whether they count as moral in the first place.

Sinnott-Armstrong seems right in holding that moral judgments cannot be delimited from other judgments simply by appeal to their content. It seems quite possible for someone to have been raised in such a way as to hold that it is morally wrong for adult men to wear shorts. And it also seems plausible that, as he also argues, moral judgments cannot be identified by reference to any sort of neurological feature common and peculiar to them and them alone. A third strategy might be to claim that moral judgments are those one makes as a result of having been inducted into a social practice that has a certain function. However, this function cannot simply be to help facilitate the sorts of social interactions that enable societies to flourish and persist, since too many obviously non-moral judgments do this.

Beyond the problem just described, attempts to pick out moral codes in the descriptive sense by appeal to their function often seem to be specifying the function that the theorist thinks morality, in the normative sense, would serve, rather than the function that actual moralities do serve. For example, Greene claims that

morality is a set of psychological adaptations that allow otherwise selfish individuals to reap the benefits of cooperation, (2013: 23)

and Haidt claims that

moral systems are interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate self-interest and make cooperative societies possible. (2011: 270)

But these claims need to deal with the existence of dysfunctional moralities that do not in fact serve these functions. Perhaps this problem could be alleviated by pointing out that many instances of a kind that have a function—for example, an actual human heart—fail to fulfill that function.

Even if Sinnott-Armstrong's position is correct with regard to morality in the descriptive sense, there might nevertheless be a code of conduct that, given certain specified conditions, would be put forward by all rational agents. That is, even if the descriptive sense of morality is a family-resemblance notion, vaguely bordered and open-textured, or even if it is significantly disjunctive and disunified, the normative sense might not be. By way of comparison, we might think of the notion of food in two ways: as what people regard as food, and as what they would regard as food if they were rational and fully informed. Certainly there is not much that unifies the first category: not even being digestible or nutritious, since people regard various indigestible and non-nutritious substances as food, and forego much that is digestible and nutritious. But that does not mean that we cannot theorize about what it would be rational to regard as food.

2. Descriptive Definitions of “Morality”

An initial naïve attempt at a descriptive definition of “morality” might take it to refer to the most important code of conduct put forward by a society

and accepted by the members of that society. But the existence of large and heterogeneous societies raises conceptual problems for such a descriptive definition, since there may not be any such society-wide code that is regarded as most important. As a result, a definition might be offered in which “morality” refers to the most important code of conduct put forward and accepted by any group, or even by an individual. Apart from containing some prohibitions on harming (certain) others, different moralities—when “morality” is understood in this way—can vary in content quite substantially.

Etiquette is sometimes included as a part of morality, applying to norms that are considered less serious than the kinds of norms for behavior that are more central to morality. Hobbes expresses this sort of view when he uses the term “small morals” to describe “decency of behavior, as how one man should salute another, or how a man should wash his mouth or pick his teeth before company”, and distinguishes these from “those qualities of mankind that concern their living together in peace and unity” (1660 [1994]: Chapter XI, paragraph 1). When etiquette is included as part of morality, morality is almost always being understood in the descriptive sense. One reason for this is that it is clear that the rules of etiquette are relative to a society or group. Moreover, there are no plausible conditions under which we could pick out the “correct” rules of etiquette as those that would be accepted by all rational beings.

Law is distinguished from morality by having explicit written rules, penalties, and officials who interpret the laws and apply the penalties. Although there is often considerable overlap in the conduct governed by morality and that governed by law, laws are often evaluated—and changed—on moral grounds. Some theorists, including Ronald Dworkin (1986), have even maintained that the interpretation of law must make use of morality.

Although the morality of a group or society may derive from its religion, morality and religion are not the same thing, even in that case. Morality is only a guide to conduct, whereas religion is always more than this. For example, religion includes stories about events in the past, usually about supernatural beings, that are used to explain or justify the behavior that it prohibits or requires. Although there is often a considerable overlap in the conduct prohibited or required by religion and that prohibited or required by morality, religions may prohibit or require more than is prohibited or required by guides to behavior that are explicitly labeled as moral guides, and may recommend some behavior that is prohibited by morality. Even when morality is not regarded as the code of conduct that is put forward by a formal religion, it is often thought to require some religious explanation and justification. However, just as with law, some religious practices and precepts are criticized on moral grounds, e.g., that the practice or precept involves discrimination on the basis of race, gender, or sexual orientation.

When “morality” is used simply to refer to a code of conduct put forward by an actual group, including a society, even if it is distinguished from etiquette, law, and religion, it is being used in a descriptive sense. It is also being used in the descriptive sense when it refers to important attitudes of individuals. Just as one can refer to the morality of the Greeks, so one can refer to the morality of a particular person. This descriptive use of “morality” is now becoming more prominent because of the work of psychologists such as Jonathan Haidt (2006), who have been influenced by the views of David Hume (1751), including his attempt to present a naturalistic account of moral judgments.

Guides to behavior that are regarded as moralities normally involve avoiding and preventing harm to others (Frankena 1980), and perhaps some norm of honesty (Strawson 1961). But all of them involve other matters as well, and Hare’s view of morality as that which is most

important allows that these other matters may be more important than avoiding and preventing harm to others (Hare 1952, 1963, 1981). This view of morality as concerning that which is most important to a person or group allows matters related to religious practices and precepts, or matters related to customs and traditions, e.g., purity and sanctity, to be more important than avoiding and preventing harm.

When “morality” is used in a descriptive sense, moralities can differ from each other quite extensively in their content and in the foundation that members of the society claim their morality to have. Some societies may claim that their morality, which is more concerned with purity and sanctity, is based on the commands of God. The descriptive sense of “morality”, which allows for the view that morality is based on religion in this way, picks out codes of conduct that are often in significant conflict with all normative accounts of morality.

A society might have a morality that takes accepting its traditions and customs, including accepting the authority of certain people and emphasizing loyalty to the group, as more important than avoiding and preventing harm. Such a morality might not count as immoral any behavior that shows loyalty to the preferred group, even if that behavior causes significant harm to innocent people who are not in that group. The familiarity of this kind of morality, which makes in-group loyalty almost equivalent to morality, seems to allow some comparative and evolutionary psychologists, including Frans De Waal (1996), to regard non-human animals to be acting in ways very similar to those that are regarded as moral.

Although all societies include more than just a concern for minimizing harm to (some) human beings in their moralities, this feature of morality, unlike purity and sanctity, or accepting authority and emphasizing loyalty, is included in everything that is regarded as a morality by any society.

Because minimizing harm can conflict with accepting authority and emphasizing loyalty, there can be fundamental disagreements within a society about the morally right way to behave in particular kinds of situations. Philosophers such as Bentham (1789) and Mill (1861), who accept a normative account of morality that takes the avoiding and preventing harm element of morality to be most important, criticize all actual moralities (referred to by “morality” in the descriptive sense) that give precedence to purity and loyalty when they are in conflict with avoiding and preventing harm.

Some psychologists, such as Haidt, take morality to include concern with, at least, all three of the triad of (1) harm, (2) purity, and (3) loyalty, and hold that different members of a society can and do take different features of morality to be most important. But beyond a concern with avoiding and preventing such harms to members of certain groups, there may be no common content shared by all moralities in the descriptive sense. Nor may there be any common justification that those who accept morality claim for it; some may appeal to religion, others to tradition, and others to rational human nature. Beyond the concern with harm, the only other feature that all descriptive moralities have in common is that they are put forward by an individual or a group, usually a society, in which case they provide a guide for the behavior of the people in that group or society. In the descriptive sense of “morality”, morality may not even incorporate impartiality with regard to all moral agents, and it may not be universalizable in any significant way (compare MacIntyre 1957).

Although most philosophers do not use “morality” in any of the above descriptive senses, some philosophers do. Ethical relativists such as Harman (1975), Westermarck (1960), and Prinz (2007), deny that there is any universal normative morality and claim that the actual moralities of societies or individuals are the only moralities there are. These relativists hold that only when the term “morality” is used in this descriptive sense is

there something that “morality” actually refers to. They claim that it is a mistake to take “morality” to refer to a universal code of conduct that, under certain conditions, would be endorsed by all rational persons. Although ethical relativists admit that many speakers of English use “morality” to refer to such a universal code of conduct, they claim such persons are mistaken in thinking that there is anything that is the referent of the word “morality” taken in that sense.

Wong (1984, 2006, 2014) claims to be an ethical relativist because he denies that there is any universal moral code that would be endorsed by all rational people. But what seems to stand behind this claim is the idea that there are cultural variations in the relative weights given to, for example, considerations of justice and considerations of interpersonal responsibility. And he assumes that those who believe in a universal morality are committed to the idea that “if there is fundamental disagreement, someone has got it wrong” (2014: 339). But Gert (2005) is certainly not a relativist, and it is central to his moral theory that there are fundamental disagreements in the rankings of various harms and benefits, and with regard to who is protected by morality, and no unique right answer in such cases. Wong himself is willing to say that some moralities are better than others, because he thinks that the moral domain is delimited by a functional criterion: among the functions of a morality are that it promote and regulate social cooperation, help individuals rank their own motivations, and reduce harm.

When used with its descriptive sense, “morality” can refer to codes of conduct with widely differing content, and still be used unambiguously. This parallels the way in which “law” is used unambiguously even though different societies have laws with widely differing content. However, when “morality” is used in its descriptive sense, it sometimes does not refer to the code of a society, but to the code of a group or an individual. As a result, when the guide to conduct put forward by, for example, a

religious group conflicts with the guide to conduct put forward by a society, it is not clear whether to say that there are conflicting moralities, conflicting elements within morality, or that the code of the religious group conflicts with morality.

In small homogeneous societies there may be a guide to behavior that is put forward by the society and that is accepted by (almost) all members of the society. For such societies there is (almost) no ambiguity about which guide “morality” refers to. However, in larger societies people often belong to groups that put forward guides to behavior that conflict with the guide put forward by their society, and members of the society do not always accept the guide put forward by their society. If they accept the conflicting guide of some other group to which they belong (often a religious group) rather than the guide put forward by their society, in cases of conflict they will regard those who follow the guide put forward by their society as acting immorally.

In the descriptive sense of “morality”, a person’s own morality cannot be a guide to behavior that that person would prefer others not to follow. However, that fact that an individual adopts a moral code of conduct for his own use does not entail that the person *requires* it to be adopted by anyone else. An individual may adopt for himself a very demanding moral guide that he thinks may be too difficult for most others to follow. He may judge people who do not adopt his code of conduct as not being as morally good as he is, without judging them to be immoral if they do not adopt it. However, such cases do not undermine the restriction; a guide is plausibly referred to as a morality only when the individual would be *willing* for others to follow it, at least if “follow” is taken to mean “successfully follow”. For it may be that the individual would not be willing for others to *try* to follow that code, because of worries about the bad effects of predictable failures due to partiality or lack of sufficient foresight or intelligence.

3. Implicit and Explicit Definitions in Allied Fields

Philosophers, because they do not need to produce operational tests or criteria in the way that psychologists, biologists, and anthropologists do, often simply take for granted that everyone knows what belongs, and does not belong, to the moral domain. This attitude finds expression in the philosopher's common appeal to intuition, or to what everyone agrees about. For example, Michael Smith (1994) provides a very detailed analysis of normative reasons, but in distinguishing specifically moral reasons from other sorts of reasons, he says only that they are picked out by appeal to a number of platitudes. And he makes no effort to provide anything like a comprehensive list of such platitudes. Moreover, it is very likely that there will be disagreement as to what counts as platitudinous. Or, if it is definitional of "platitude" that it be uncontroversial, it may be that what is platitudinous about morality will be so thin as to fail to separate morality from other domains. Failing to specify which particular criteria one takes to govern one's own theorizing, and consequently tacitly relying on the idea that everyone already knows what counts as moral, can lead to a number of problems. One, of course, is a conflation of morality with other things (see Machery 2012 on Churchland 2011). Another is that one mistakes one's own cultural biases for universal truths (Haidt and Kesiber 2010).

Because theorists in psychology and anthropology often need to design questionnaires and other sorts of probes of the attitudes of subjects, they might be expected to be more sensitive to the need for a reasonably clear means of separating moral judgments from other sorts of judgments. After all, examining the specifically moral judgments of individuals is one of the most direct means of determining what the moral code of a person or group might be. But despite this expectation, and roughly half a century ago, Abraham Edel (1962: 56) decried the lack of an explicit concern to delimit the domain of morality among anthropologists, writing that

"morality...is taken for granted, in the sense that one can invoke it or refer to it at will; but it is not explained, depicted, or analysed". One explanation for this that Edel suggested is the same as the explanation for the same phenomenon in Philosophy: "it is assumed that we all know what morality is and no explicit account need be given". But the danger for those making this assumption, he points out, is that of "merging the morality concept with social control concepts". Reinforcing this tendency was the influence, in anthropology, of the sociologist Émile Durkheim (1906 [2009]), for whom morality was simply a matter of how a given society enforces whatever social rules it happens to have.

The failure to offer an operational definition of morality or moral judgment may help explain the widespread but dubious assumption in contemporary anthropology, noted by James Laidlaw (2016: 456), that altruism is the essential and irreducible core of ethics. But Laidlaw also notes that many of the features of what Bernard Williams (1985) described as "the morality system"—features that Williams himself criticized as the parochial result of a secularization of Christian values—are in fact widely shared outside of the West. This state of affairs leads Laidlaw to ask the crucial question:

Which features, formal or substantive, are shared by the "morality system" of the modern West and those of the other major agrarian civilizations and literate religions?

This is, to a very close approximation, a request for the definition of morality in the descriptive sense.

Klenk (2019) notes that in recent years anthropology has taken what he terms an "ethical turn", recognizing moral systems, and ethics more generally, as a distinct object of anthropological study. This is a move away from the Durkheimian paradigm, and includes the study of self-

development, virtues, habits, and the role of explicit deliberation when moral breakdowns occur. However, Klenk's survey of attempts by anthropologists to study morality as an independent domain lead him to conclude that, so far, their efforts do

not readily allow a distinction between moral considerations and other normative considerations such as prudential, epistemic, or aesthetic ones. (2019: 342)

In light of Edel's worry about a conflation of moral systems with systems of social control, it is interesting to consider Curry (2016), who defends the hypothesis that

morality turns out to be a collection of biological and cultural solutions to the problems of cooperation and conflict recurrent in human social life. (2016: 29)

Curry notes that rules related to kinship, mutualism, exchange, and various forms of conflict resolution appear in virtually all societies. And he argues that many of them have precursors in animal behavior, and can be explained by appeal to his central hypothesis of morality as a solution to problems of cooperation and conflict resolution. He also notes that philosophers, from Aristotle through Hume, Russell, and Rawls, all took cooperation and conflict resolution to be central ideas in understanding morality. It is unclear, however, whether Curry's view can adequately distinguish morality from law and from other systems that aim to reduce conflict by providing solutions to coordination problems.

Turning from anthropology to psychology, one significant topic of investigation is the existence and nature of a distinction between the moral and the conventional. More specifically, the distinction at issue is between (a) acts that are judged wrong only because of a contingent convention or because they go against the dictates of some relevant authority, and (b)

those that are judged to be wrong quite independently of these things, that have a seriousness to them, and that are justified by appeal to the notions of harm, rights, or justice. Elliot Turiel emphasized this distinction, and drew attention to the danger, if one overlooks it, of lumping together moral rules with non-moral "conventions that further the coordination of social interactions within social systems" (1983: 109–111). Those who accept this distinction are implicitly offering a definition of morality in the descriptive sense. Not everyone does accept the distinction, however. Edouard Machery and Ron Mallon (2010) for example, are suspicious of the idea that authority-independence, universality, justification by appeal to harm, justice, or rights, and seriousness form a cluster found together with sufficient regularity to be used to set moral norms apart from other norms. Kelly et al. (2007) are similarly skeptical, and bring empirical evidence to bear on the question.

The psychologist Kurt Gray might be seen as offering an account of moral judgment that would allow us to determine the morality of an individual or group. He and his co-authors suggest that

morality is essentially represented by a cognitive template that combines a perceived intentional agent with a perceived suffering patient. (Gray, Young, & Waytz 2012: 102)

This claim, while quite strong, is nevertheless not as implausibly strong as it might seem, since the thesis is directly concerned with the *template* we use when thinking about moral matters; it is not directly concerned with the nature of morality itself. In the sense of "template" at issue here, the template we use when thinking about dogs might include having four legs, a tail, and fur, among other things. But that does not mean that an animal must have these features to count as a dog, or even that we believe this.

Given the way that Gray et al. think of templates, even if their hypothesis is correct, it would not mean that our psychology requires us to think of the moral as always involving intentional agents and perceiving patients. In line with this, and despite some lapses in which they suggest that “moral acts can be *defined* in terms of intention and suffering”, (2012: 109) their considered view seems to be only that the dyadic template fits the *majority* of moral situations, as we conceive them. Moreover, the link between immoral behavior and suffering to which they appeal in defending their general view is sometimes so indirect as to undermine its significance. For example, they fit authority violations into their suffering-based template by noting that “authority structures provide a way of peacefully resolving conflict” and that “violence results when social structures are threatened”. In a similar stretch, they account for judgments that promiscuity is wrong by gesturing at the suffering involved in sexually transmitted diseases (2012: 107).

Another position in cognitive psychology that has relevance for the definition of morality in the descriptive sense takes moral judgment to be a natural kind: the product of an innate moral grammar (Mikhail 2007). If moral judgment is a natural kind in this way, then a person’s moral code might simply consist in the moral judgments that person is disposed to make. One piece of evidence that there is such a grammar is to be found in the relative universality of certain moral concepts in human cultures: concepts such as obligation, permission, and prohibition. Another is an argument similar to Chomsky’s famous “poverty of the stimulus” argument for a universal human grammar (Dwyer et al. 2010; see also Roedder and Harman 2010).

In evolutionary biology, morality is sometimes simply equated with fairness (Baumard et al. 2013: 60, 77) or reciprocal altruism (Alexander 1987: 77). But it is also sometimes identified by reference to an evolved capacity to make a certain sort of judgment and perhaps also to signal that

one has made it (Hauser 2006). This also makes morality into something very much like a natural kind, that can be identified by reference to causal/historical processes. In that case, a content-based definition of morality isn’t required: certain central features are all that one needs to begin one’s theorizing, since they will be enough to draw attention to certain psychologically and biologically individuated mechanisms, and the study of morality will be a detailed inquiry into the nature and evolutionary history of these mechanisms.

4. Normative Definitions of “Morality”

Those who use “morality” normatively hold that morality is (or would be) the behavioral code that meets the following condition: all rational persons, under certain specified conditions, would endorse it. Indeed, this is a plausible basic schema for definitions of “morality” in the normative sense. Although some hold that no code could meet the condition, many theorists hold that there is one that does; we can call the former “moral skeptics” and the latter “moral realists” (see entries on LINK: moral skepticism and moral realism).

Many moral skeptics would reject the claim that there are any universal ethical truths, where the ethical is a broader category than the moral. But another interesting class of moral skeptics includes those who think that we should only abandon the narrower category of the moral—partly because of the notion of *a code* that is central to that category. These moral skeptics hold that we should do our ethical theorizing in terms of the good life, or the virtues. Elizabeth Anscombe (1958) gave expression to this kind of view, which also finds echoes in the work of Bernard Williams (1985). On the other hand, some virtue theorists might take perfect rationality to entail virtue, and might understand morality to be something like the code that such a person would implicitly endorse by

acting in virtuous ways. In that case, even a virtue theorist might count as a moral realist in the sense above.

Consequentialist views might not seem to fit the basic schema for definitions of “morality” in the normative sense, since they do not appear to make reference to the notions of endorsement or rationality. But this appearance is deceptive. Mill himself explicitly defines morality as

the rules and precepts for human conduct, by the observance of which [a happy existence] might be, to the greatest extent possible, secured. (1861 [2002: 12])

And he thinks that the mind is not in a “right state” unless it is in “the state most conducive to the general happiness”—in which case it would certainly favor morality as just characterized. And the act-consequentialist J.J.C. Smart (1956) is also explicit that he is thinking of ethics as the study of how it is most rational to behave. His embrace of utilitarianism is the result of his belief that maximizing utility is always the rational thing to do. On reflection it is not surprising that many moral theorists implicitly hold that the codes they offer would be endorsed by all rational people, at least under certain conditions. Unless one holds this, one will have to admit that, having been shown that a certain behavior is morally required, a rational person might simply shrug and ask “So what? What is that to me?” And, though some exceptions are mentioned below, very few moral realists think that their arguments leave this option open. Even fewer think this option remains open if we are allowed to add some additional conditions beyond mere rationality: a restriction on beliefs, for example (similar to Rawls’ (1971: 118) veil of ignorance), or impartiality.

Definitions of morality in the normative sense—and, consequently, moral theories—differ in their accounts of rationality, and in their specifications of the conditions under which all rational persons would necessarily

endorse the code of conduct that therefore would count as morality. These definitions and theories also differ in how they understand what it is to endorse a code in the relevant way. Related to these differences, definitions of “morality”—and moral theories—differ with regard to those to whom morality applies: that is, those whose behavior is subject to moral judgment. Some hold that morality applies only to those rational beings that have certain specific features of human beings: features that make it rational for them to endorse morality. These features might, for example, include fallibility and vulnerability. Other moral theories claim to put forward an account of morality that provides a guide to all rational beings, even if these beings do not have these human characteristics, e.g., God.

Among those who use “morality” normatively, virtually all hold that “morality” refers to a code of conduct that *applies to* all who can understand it and can govern their behavior by it, though many hold that it *protects* a larger group. Among such theorists it is also common to hold that morality should never be overridden. That is, it is common to hold that no one should ever violate a moral prohibition or requirement for non-moral reasons. This claim is trivial if “should” is taken to mean “morally should”. So the claim about moral overridingness is typically understood with “should” meaning “rationally should”, with the result that moral requirements are asserted to be rational requirements. Though common, this view is by no means always taken as definitional. Sidgwick (1874) despaired of showing that rationality *required* us to choose morality over egoism, though he certainly did not think rationality required egoism either. More explicitly, Gert (2005) held that though moral behavior is always rationally *permissible*, it is not always rationally *required*. Foot (1972) seems to have held that any reason—and therefore any rational requirement—to act morally would have to stem from a contingent commitment or an objective interest. And she also seems to have held that sometimes neither of these sorts of reasons might be available, so that moral behavior might not be rationally required for some agents. Finally,

moral realists who hold desire-based theories of reasons and formal, means/end theories of rationality sometimes explicitly deny that moral behavior is always even rationally *permissible* (Goldman 2009), and in fact this seems to be a consequence of Foot's view as well, though she does not emphasize it.

Despite the fact that theorists such as Sidgwick, Gert, Foot, and Goldman do not hold that moral behavior is rationally required, they are by no means precluded from using "morality" in the normative sense. Using "morality" in the normative sense, and holding that there is such a thing, only entails holding that rational people would put a certain system forward; it does not entail holding that rational people would always be motivated to follow that system themselves. But to the degree that a theorist would deny even the claim about endorsement, and hold instead that rational people might not only fail to act morally, but might even reject it as a public system, that theorist is either not using "morality" in a normative sense, or is denying the existence of morality in that sense. Such a theorist may also be using "morality" in a descriptive sense, or may not have any particular sense in mind.

When "morality" is used in its normative sense, it need not have either of the two formal features that are essential to moralities referred to by the descriptive sense: that it be a code of conduct that is put forward by a society, group, or individual, or that it be accepted as a guide to behavior by the members of that society or group, or by that individual. Indeed, it is possible that morality, in the normative sense, has never been put forward by any particular society, by any group at all, or even by any individual. This is partly a consequence of the fact that "morality" in the normative sense is understood in terms of a conditional that is likely to be counterfactual: it is the code that would be endorsed by any fully rational person under certain conditions.

If one is a moral realist, and one also acknowledges the descriptive sense of "morality", one may require that descriptive moralities at least approximate, in some ways, morality in the normative sense. That is, one might claim that the guides to behavior of some societies lack so many of the essential features of morality in the normative sense, that it is incorrect to say that these societies even have a morality in a descriptive sense. This is an extreme view, however. A more moderate position would hold that all societies have something that can be regarded as their morality, but that many of these moralities—perhaps, indeed, all of them—are defective. That is, a moral realist might hold that although these actual guides to behavior have enough of the features of normative morality to be classified as descriptive moralities, they would not be endorsed in their entirety by all moral agents.

While moral realists do not claim that any actual society has or has ever had morality as its actual guide to conduct, "natural law" theories of morality claim that any rational person in any society, even one that has a defective morality, is capable of knowing what general kinds of actions morality prohibits, requires, discourages, encourages, and allows. In the theological version of natural law theories, such as that put forward by Aquinas, this is because God implanted this knowledge in the reason of all persons. In the secular version of natural law theories, such as that put forward by Hobbes (1660), natural reason is sufficient to allow all rational persons to know what morality prohibits, requires, etc. Natural law theorists also claim that morality applies to all rational persons, not only those now living, but also those who lived in the past.

In contrast to natural law theories, other moral theories do not hold quite so strong a view about the universality of knowledge of morality. Still, many hold that morality is known to all who can legitimately be judged by it. Baier (1958), Rawls (1971) and contractarians deny that there can be an *esoteric* morality: one that judges people even though they cannot know

what it prohibits, requires, etc. For all of the above theorists, morality is what we can call a *public system*: a system of norms (1) that is knowable by all those to whom it applies and (2) that is not irrational for any of those to whom it applies to follow (Gert 2005: 10). Moral judgments of blame thus differ from legal or religious judgments of blame in that they cannot be made about persons who are legitimately ignorant of what they are required to do. Act consequentialists seem to hold that everyone should know that they are morally required to act so as to bring about the best consequences, but even they do not seem to think judgments of moral blame are appropriate if a person is legitimately ignorant of what action would bring about the best consequences (Singer 1993: 228). Parallel views seem to be held by rule consequentialists (Hooker 2001: 72).

The ideal situation for a legal system would be that it be a public system. But in any large society this is not possible. Games are closer to being public systems and most adults playing a game know its rules, or they know that there are judges whose interpretation determines what behavior the game prohibits, requires, etc. Although a game is often a public system, its rules apply only to those playing the game. If a person does not care enough about the game to abide by the rules, she can usually quit. Morality is the one public system that no rational person can quit. The fact that one cannot quit morality means that one can do nothing to escape being legitimately liable to sanction for violating its norms, except by ceasing to be a moral agent. Morality applies to people simply by virtue of their being rational persons who know what morality prohibits, requires, etc., and being able to guide their behavior accordingly.

Public systems can be *formal* or *informal*. To say a public system is informal is to say that it has no authoritative judges and no decision procedure that provides a unique guide to action in all situations, or that resolves all disagreements. To say that a public system is formal is to say that it has one or both of these things (Gert 2005: 9). Professional

basketball is a formal public system; all the players know that what the referees call a foul determines what is a foul. Pickup basketball is an informal public system. The existence of persistent moral disagreements shows that morality is most plausibly regarded as an informal public system. This is true even for such moral theories as the Divine Command theory and act utilitarianism, inasmuch as there are no authoritative judges of God's will, or of which act will maximize utility, and there are no decision procedures for determining these things (Scanlon 2011: 261–2). When persistent moral disagreement is recognized, those who understand that morality is an informal public system admit that how one should act is morally unresolvable, and if some resolution is required, the political or legal system can be used to resolve it. These formal systems have the means to provide unique guides, but they do not provide the uniquely correct moral guide to the action that should be performed.

An important example of a moral problem left unsettled by the informal public system of morality is whether fetuses are impartially protected by morality and so whether or under what conditions abortions are allowed. There is continuing disagreement among fully informed moral agents about this moral question, even though the legal and political system in the United States has provided fairly clear guidelines about the conditions under which abortion is legally allowed. Despite this important and controversial issue, morality, like all informal public systems, presupposes agreement on how to act in most moral situations, e.g., all agree that killing or seriously harming any moral agent requires strong justification in order to be morally allowed. No one thinks it is morally justified to cheat, deceive, injure, or kill a moral agent simply in order to gain sufficient money to take a fantastic vacation. Moral matters are often thought to be controversial because everyday decisions, about which there is no controversy, are rarely discussed. The amount of agreement concerning what rules are moral rules, and on when it is justified to violate

one of these rules, explains why morality can be a public system even though it is an informal system.

By using the notion of an informal public system, we can improve the basic schema for definitions of “morality” in the normative sense. The old schema was that morality is the *code* that all rational persons, under certain specified conditions, would endorse. The improved schema is that morality is the *informal public system* that all rational persons, under certain specified conditions, would endorse. Some theorists might not regard the informal nature of the moral system as definitional, holding that morality might give knowable precise answers to every question. This would have the result that conscientious moral agents often cannot know what morality permits, requires, or allows. Some philosophers deny that this is a genuine possibility.

On any definition of “morality”, whether descriptive or normative, it is a code of conduct. However, on ethical- or group-relativist accounts or on individualistic accounts—all of which are best regarded as accounts of morality in the descriptive sense—morality often has no special content that distinguishes it from nonmoral codes of conduct, such as law or religion. Just as a legal code of conduct can have almost any content, as long as it is capable of guiding behavior, and a religious code of conduct has no limits on content, most relativist and individualist accounts of morality place few limits on the content of a moral code. Of course, actual codes do have certain minimal limits—otherwise the societies they characterize would lack the minimum required degree of social cooperation required to sustain their existence over time. On the other hand, for moral realists who explicitly hold that morality is an informal public system that all rational persons would put forward for governing the behavior of all moral agents, it has a fairly definite content. Hobbes (1660), Mill (1861), and most other non-religiously influenced

philosophers in the Anglo-American tradition limit morality to behavior that, directly or indirectly, affects others.

The claim that morality only governs behavior that affects others is somewhat controversial, and so probably should not be counted as definitional of morality, even if it turns out to be entailed by the correct moral theory. Some have claimed that morality also governs behavior that affects only the agent herself, such as taking recreational drugs, masturbation, and not developing one’s talents. Kant (1785) may provide an account of this wide concept of morality. Interpreted this way, Kant’s theory still fits the basic schema, but includes these self-regarding moral requirements because of the particular account of rationality he employs. However, *pace* Kant, it is doubtful that all moral agents would put forward a universal guide to behavior that governs behavior that does not affect them at all. Indeed, when the concept of morality is completely distinguished from religion, moral rules do seem to limit their content to behavior that directly or indirectly causes or risks harm to others. Some behavior that seems to affect only oneself, e.g., taking recreational drugs, may have a significant indirect harmful effect on others by supporting the illegal and harmful activity of those who benefit from the sale of those drugs.

Confusion about the content of morality sometimes arises because morality is not distinguished sufficiently from religion. Regarding self-affecting behavior as governed by morality is supported by the idea that we are created by God and are obliged to obey God’s commands, and so may be a holdover from the time when morality was not clearly distinguished from religion. This religious holdover might also affect the claim that some sexual practices such as homosexuality are immoral. Those who clearly distinguish morality from religion typically do not regard sexual orientation as a moral matter.

It is possible to hold that having a certain sort of social goal is definitional of morality (Frankena 1963). Stephen Toulmin (1950) took it to be the harmony of society. Baier (1958) took it to be “the good of everyone alike”. Utilitarians sometimes claim it is the production of the greatest good. Gert (2005) took it to be the lessening of evil or harm. This latter goal may seem to be a significant narrowing of the utilitarian claim, but utilitarians always include the lessening of harm as essential to producing the greatest good and almost all of their examples involve the avoiding or preventing of harm. It is notable that the paradigm cases of moral rules are those that prohibit causing harm directly or indirectly, such as rules prohibiting killing, causing pain, deceiving, and breaking promises. Even those precepts that require or encourage positive action, such as helping the needy, are almost always related to preventing or relieving harms, rather than promoting goods such as pleasure.

Among the views of moral realists, differences in content are less significant than similarities. For all such philosophers, morality prohibits actions such as killing, causing pain, deceiving, and breaking promises. For some, morality also requires charitable actions, but failure to act charitably on every possible occasion does not require justification in the same way that any act of killing, causing pain, deceiving, and breaking promises requires justification. Both Kant (1785) and Mill (1861) distinguish between duties of perfect obligation and duties of imperfect obligation and regard not harming as the former kind of duty and helping as the latter kind of duty. For Gert (2005), morality encourages charitable action, but does not require it; it is always morally good to be charitable, but it is not immoral not to be charitable.

Even if the plausible basic schema for definitions of “morality” in the normative sense is accepted, one’s understanding of what morality is, in this sense, will still depend very significantly on how one understands rationality. As has already been mentioned, morality, in the normative

sense, is sometimes taken to prohibit certain forms of consensual sexual activity, or the use of recreational drugs. But including such prohibitions in an account of morality as a universal guide that all rational persons would put forward requires a very particular view of rationality. After all, many will deny that it is irrational to favor harmless consensual sexual activities, or to favor the use of certain drugs for purely recreational purposes.

One concept of rationality that supports the exclusion of sexual matters, at least at the basic level, from the norms of morality, is that for an action to count as irrational it must be an act that harms oneself without producing a compensating benefit for someone—perhaps oneself, perhaps someone else. Such an account of rationality might be called “hybrid”, since it gives different roles to self-interest and to altruism. An account of morality based on the hybrid concept of rationality could agree with Hobbes (1660) that morality is concerned with promoting people living together in peace and harmony, which includes obeying the rules prohibiting causing harm to others. Although moral prohibitions against actions that cause harm or significantly increase the risk of harm are not absolute, in order to avoid acting immorally, justification is always needed when violating these prohibitions. Kant (1797) seems to hold that it is never justified to violate some of these prohibitions, e.g., the prohibition against lying. This is largely a result of the fact that Kant’s (1785) concept of rationality is purely formal, in contrast with the hybrid concept of rationality described above.

Most moral realists who offer moral theories do not bother to offer anything like a definition of morality. Instead, what these philosophers offer is a theory of the nature and justification of a set of norms with which they take their audience already to be acquainted. In effect, they tacitly pick morality out by reference to certain salient and relative uncontroversial bits of its content: that it prohibits killing, stealing,

deceiving, cheating, and so on. In fact, this would not be a bad way of defining morality, if the point of such a definition were only to be relatively theory-neutral, and to allow theorizing to begin. We could call it “the reference-fixing definition” or “the substantive definition” (see Prinz and Nichols 2010: 122).

Some, including Hare (1952, 1963), have been tempted to argue against the possibility of a substantive definition of morality, on the basis of the claim that moral disapproval is an attitude that can be directed at anything whatsoever. Foot (1958a, 1958b), argued against this idea, but the substantive definition still has the drawback is that it does not, somehow, seem to get at the essence of morality. One might suggest that the substantive definition has the advantage of including Divine Command theories of morality, while such theories might seem to make trouble for definitions based on the plausible schema given above. But it is plausible to hold that Divine Command theories rest on Natural Law theories, which do in fact fit the schema. Divine Command theories that do not rest on Natural Law might make trouble for the schema, but one might also think that such theories rest instead on a confusion, since they seem to entail that God might have made it immoral to act beneficently.

5. Variations

As one gives more substance and detail to the general notions of endorsement, rationality, and the relevant conditions under which rational people would endorse morality, one moves further from providing a definition of morality in the normative sense, and closer to providing an actual moral theory. And a similar claim is true for definitions of morality in the descriptive sense, as one specifies in more detail what one means in claiming that a person or group endorses a system or code. In the following four subsections, four broad ways of making the definitions of

morality more precise are presented. They are all sufficiently schematic to be regarded as varieties of definition, rather than as theories.

5.1 Morality as linked to norms for responses to behavior

Expressivists about morality do not take there to be any objective content to morality that could underwrite what we above called “the substantive definition”. Rather, they explicitly recognize the existence of significant variation in what rules and ideals different people put forward as morality in the normative sense. And they doubt that this variation is compatible with moral realism. Consequently, they need to offer some unifying features of these different sets of rules and ideals, despite variation in their content. As a result of this pressure, some expressivists end up offering explicit accounts of a *distinctively* moral attitude one might hold towards an act token or type. These accounts can of course be taken to underwrite various forms of morality in the descriptive sense. But they can also be taken to provide the basis of one form of moral realism.

To see how an expressivist view can be co-opted by a moral realist of a certain sort, consider Allan Gibbard’s (1990) moral expressivism. Gibbard holds that moral judgments are expressions of the acceptance of norms for feeling the emotions of guilt and anger. One can accept Gibbard’s view of what it is to endorse a moral claim without accepting the view that, in conflicts, all disagreements are faultless. That is, even a moral realist can use Gibbard’s view of the nature of moral judgment, and extract from it a definition of morality. Used by such a theorist, Gibbard’s view entails that morality, in the normative sense, is the code that is picked out by the *correct* set of norms for feeling guilt and anger: that is, the norms a *rational* person would endorse. This is equivalent to accepting the plausible general schema for a definition of “morality” given above, and understanding endorsement in a special sense. To endorse a code in the

relevant way, on this definition, is to think that violations of its norms make guilt and anger appropriate.

Closely related to Gibbard's account is one according to which the norms of relevance are not norms for the emotions, but are norms for other reactions to behavior. For example, a person's morality might be the set of rules and ideals they regard as picked out by appropriate norms for praise and blame, and other social sanctions (Sprigge 1964: 317). In fact, reference to praise and blame may be more adequate than reference to guilt and anger, since the latter seem only to pick out moral prohibitions, and not to make room for the idea that morality also recommends or encourages certain behaviors even if it does not require them. For example, it is plausible that there is such a thing as supererogatory action, and that the specification of what counts as supererogatory is part of morality—whether in the descriptive or normative sense. But it does not seem likely that we can account for this part of morality by appeal to norms for guilt and anger, and it is not at all clear that there are emotions that are as closely linked to supererogation as guilt and anger are to moral transgression. On the other hand, it seems plausible that norms for praising action might help to pick out what counts as supererogatory.

Another version of the present strategy would replace talk of praise and blame with talk of reward and punishment. This view would take morality to be a system that explained what kinds of actions are appropriately rewarded and—more centrally—punished. This sort of view, which remains closely related to Gibbard's suggestion, can also be regarded as fitting the general schema given above. On this view, the notion of endorsing a code is unpacked in terms of the acceptance of norms for reward and punishment. Skorupski (1993), following Mill (1861), advocates a definition of morality along these lines, though he then understands punishment primarily in terms of blame, and understands blame as very closely linked to emotion—indeed, merely having the

emotion can count as blaming—so that the resulting view is similar to Gibbard's in one important way, at least when one focuses on moral wrongness.

It is certainly plausible that it is appropriate to feel guilt when one acts immorally, and to feel anger at those who act immorally towards those one cares about. It is even plausible that it is *only* appropriate, in some particular sense of “appropriate”, to feel guilt and anger in connection with moral transgressions. So norms for guilt and anger may well uniquely pick out certain moral norms. And similar claims might be made about norms for praise and blame. However, it is not equally clear that morality is properly *defined* in terms of emotions or other reactions to behavior. For it may be, as Skorupski emphasizes, that we need to understand guilt and anger, and praise and blame, in terms of moral concepts. This worry about direction of explanation seems less pressing for the notions of reward and punishment. These responses to behavior, at least *in themselves*, might simply be understood in terms of the meting out of benefits and harms. Of course they will only *count as reward and punishment* when they are linked to someone's having followed or violated a rule that all rational people would want to see enforced by such responses.

5.2 Morality as linked to advocacy of a code

One way of understanding the notion of endorsement is as advocacy. Advocating a code is a second- or third-personal matter, since one advocates a code to others. Moreover, it is consistent with advocating a code, that one does not plan on following that code oneself. Just as asserting something one believes to be false still counts as asserting it, hypocritical advocacy of a code still counts as advocacy of that code. When endorsement is understood as advocacy, it can be used in definitions of morality, in the descriptive sense, as long as it is the morality of a group or society. And advocacy can also be used as an interpretation of

endorsement when providing a definition of morality in the normative sense. Of course those who accept a definition of morality in any of these senses—as the code that a group or society endorses, or as the code that would be universally advocated by all rational agents under certain conditions—do not hold that the advocacy would necessarily, or even probably, be hypocritical. But they do hold that the important thing about a moral code—what picks it out as a moral code—is that it would be *put forward* by all the relevant agents, not that it would be *followed* by all of them. The notion of advocacy has less of a place in a descriptive account of a single person’s morality, since when someone is hypocritical we often deny that they really hold the moral view that they advocate.

Mill (1861), in addition to offering a moral theory, takes pains to explain how morality differs from other normative systems. For him, norms that simply promote utility are norms of expediency. In order to qualify as morally wrong, an act must be one that ought to be punished. Thinking that an act of a certain kind ought to be punished is a third-personal matter, so it seems plausible to put Mill’s view of what is definitional of morality into the category being discussed in this section. It is worth noting that hypocrisy is, for Mill, not only a possibility, but—given the present sorry state of moral education—virtually unavoidable. That is because being motivated to advocate punishment for a certain kind of act is quite different from being motivated to refrain from that same kind of act. Advocating punishment for a certain kind of act might be one’s utility-maximizing choice, while actually performing that kind of act (trying, of course, to avoid detection) might also be utility-maximizing. And for Mill what determines what a person will advocate, and how a person will act, are the foreseeable consequences *for that person*.

Bernard Gert’s (2005) moral view also operates with a definition of morality that understands endorsement as advocacy, in the sense of putting forward as a guide for all rational agents. Gert offers the following two

conditions as those under which all rational persons would put forward a universal guide for governing the behavior of all moral agents. The first condition is that they are seeking agreement with all other rational persons or moral agents. The second condition is that they use only those beliefs that are shared by all rational persons: for example, that they themselves are fallible and vulnerable and that all those to whom morality applies are also fallible and vulnerable. The second condition rules out both religious beliefs and scientific beliefs since there are no religious beliefs or scientific beliefs that all rational persons share. This condition is plausible because no universal guide to behavior that applies to all rational persons can be based on beliefs that some of these rational persons do not share.

5.3 Morality as linked to acceptance of a code

Another way of understanding the notion of endorsement is as acceptance. Unlike advocating a code, accepting a code is a first-personal matter. It might include intending to conform one’s own behavior to that code, feeling guilty when one does not, and so on. One cannot hypocritically accept a code. Indeed, hypocrisy is simply a matter of advocating a code one does not accept. So this notion of endorsement is available to someone who is trying to provide a definition of morality in the descriptive sense, even when considering a single person’s morality.

Paradigmatic views in the natural law tradition starting with Aquinas hold both that the laws of morality have their source in God, and that these laws constitute the principles of human practical rationality (Finnis 1980; MacIntyre 1999). Views in this tradition may be seen as using the basic schema for definitions of morality in the normative sense, understanding endorsement as acceptance. Members of this tradition typically hold that all rational persons know what kinds of actions morality prohibits, requires, discourages, encourages, and allows. It is central to Aquinas’s view that morality is known to all those whose behavior is subject to

moral judgment, even if they do not know of the revelations of Christianity. This is why Aquinas holds that knowing what morality prohibits and requires does not involve knowing why morality prohibits and requires what it does.

Those who belong to the natural law tradition also hold that reason endorses acting morally. This sort of endorsement of course has a cognitive component. But it is also motivational. Aquinas does not hold that knowledge of morality is always effective: it can be blotted out by evil persuasions or corrupt habits. But if reason is not opposed by such forces, any rational person would not only know what was prohibited and required by morality, but would follow those prohibitions and requirements. So, for natural law theorists, endorsement amounts to acceptance.

5.4 Morality as linked to justification to others

The lack of an explicit and widely accepted definition of morality may partially explain the resilience of act-consequentialist accounts of morality. Without an explicit definition, it may be easier to ignore the fact that act-consequentialist theories are not particularly concerned with interpersonal interactions, but typically apply just as well to desert island scenarios as to individuals who live in societies. In any case, it has been recognized that in order to combat consequentialism, it would be helpful to have something like a plausible definition of morality that made it clear that the subject matter of morality is something different from simply the goodness and badness of consequences. T.M. Scanlon (1982, 1998), applying this strategy, suggests that the subject matter of morality—what we are talking about, when we talk about morality—is a system of rules for the regulation of behavior that is not reasonably rejectable based on a desire for informed unforced general agreement.

Scanlon's suggestion regarding the subject matter of morality can easily be seen as an instance of the general schema given above. His "system of rules" is a specific kind of informal public system; he understands endorsement by all rational people as non-rejection by all reasonable people; and he offers a specific account of the conditions under which moral agents would reach the relevant agreement. But Scanlon also places very heavy emphasis on the fact that if he is right about the subject matter of morality, then what compliance with moral norms allows us to do is to justify our behavior to others in ways that they cannot reasonably reject. Indeed, the ability to justify ourselves to reasonable people is a primary source of moral motivation for Scanlon (see also Sprigge 1964: 319). This might seem to suggest a somewhat different definitional claim about morality: that morality consists in the most basic norms in terms of which we justify ourselves to others. But it is plausible that this purportedly definitional claim is better thought of as a corollary of Scanlon's particular version of the general schema, with endorsement understood as non-rejection. For, if morality is the system of norms that would be endorsed in this way, we can justify our actions to others by pointing out that even they, were they reasonable, would have endorsed rules that allowed our behavior.

Stephen Darwall's (2006) moral view can also be seen as flowing from a version of the general schema, and yielding claims about justifiability to others. Darwall claims that morality is a matter of equal accountability among free and rational beings. On his view, I behave morally towards you to the degree that I respect the claims you have authority to make on me. Darwall also holds that I will respect those claims if I acknowledge certain assumptions to which I am committed simply in virtue of being a rational, deliberating agent. As a result, his view is that morality—or at least the morality of obligation—is a "scheme of accountability" (a certain sort of informal public system) that all rational people will endorse. Unlike Scanlon's view, however, Darwall's view makes use of a stronger sense of

endorsement than non-rejection. Specifically, it includes the recognition of the reasons provided by the authoritative demands of other people. And that recognition is positively motivational.

Both Scanlon's and Darwall's views emphasize the social nature of morality, taken in the normative sense: Scanlon, by reference to justification to others; Darwall, by appeal to the relevance of second-personal reasons. But Darwall builds a responsiveness to second-personal reasons into the relevant notion of rationality, while Scanlon simply makes the empirical claim that many people are motivated by a desire to justify themselves to others, and notes that his definition of morality will yield rules that will allow one to do this, if one follows them. The sort of definition described in section 5.1 also makes the social nature of morality essential to it, since it centrally features the notion of a response to the behavior of others. The definitions described in sections 5.2 and 5.3 do not entail the social nature of morality, since it is possible to accept, and even to advocate, a code that concerns only self-regarding behavior. But on any plausible account of rationality a code that would be advocated by all moral agents will govern interpersonal interactions, and will include rules that prohibit causing harm without sufficient reason. Only the definition offered in section 5.3 therefore can be taken as realistically compatible with an egoistic morality.

Bibliography

- Alexander, Richard, 1987, *The Biology of Moral Systems*, New York: Routledge.
- Anscombe, G. E. M., 1958, "Modern Moral Philosophy", *Philosophy*, 33(124): 1–19. doi:10.1017/S0031819100037943
- Aquinas, Thomas, c.1270, *Summa Theologiae*, Paris.
- Baier, Kurt, 1958, *The Moral Point of View*, Ithaca, NY: Cornell University Press.

- Baumard, Nicolas, Jean-Baptiste André, and Dan Sperber, 2013, "A Mutualistic Approach to Morality: The Evolution of Fairness by Partner Choice", *Behavioral and Brain Sciences*, 36(1): 59–78. doi:10.1017/S0140525X11002202
- Bentham, Jeremy, 1789, *An Introduction to the Principles of Morals and Legislation*, New York: Prometheus Books, 1988.
- Brink, David, 1997, "Kantian Rationalism: Inescapability, Authority, and Supremacy", in *Ethics and Practical Reason*, Garrett Cullity and Berys Gaut (eds.), Oxford: Oxford University Press, pp. 255–291.
- Churchland, Patricia, 2011, *Braintrust: What Neuroscience Tells Us About Morality*, Princeton, NJ: Princeton University Press.
- Curry, Oliver Scott, 2016, "Morality as Cooperation: A Problem-Centred Approach", in *The Evolution of Morality*, Todd K. Shackelford and Randal Hansen (eds.), Cham: Springer, pp. 27–51. doi:10.1007/978-3-319-19671-8_2
- Darwall, Stephen, 2006, *The Second-person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.
- De Waal, Frans, 1996, *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*, Cambridge, MA: Harvard University Press.
- Doris, John M. and The Moral Psychology Research Group (eds.), 2010, *The Moral Psychology Handbook*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199582143.001.0001
- Durkheim, Émile, 1906 [2009], "La Détermination du fait moral". Collected in *Sociologie et Philosophie*, Paris: Félix Alcan, 1924. Translated as "The Determination of Moral Facts" in *Sociology and Philosophy*, David Pocock (ed. and trans.), 1953, Reprinted Routledge revivals. London: Routledge, 2009, pp. 16–31.
- Dworkin, Ronald, 1986, *Law's Empire*, Cambridge, MA: Belknap Press.
- Dwyer, Susan, Bryce Huebner, and Marc D. Hauser, 2010, "The Linguistic



- Analogy: Motivations, Results, and Speculations”, *Topics in Cognitive Science*, 2(3): 486–510. doi:10.1111/j.1756-8765.2009.01064.x
- Edel, Abraham, 1962, “Anthropology and Ethics in Common Focus”, *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 92(1): 55–72. doi:10.2307/2844321
- Finnis, John, 1980, *Natural Law and Natural Rights*, Oxford: Oxford University Press.
- Foot, Philippa, 1958a, “Moral Beliefs”, *Proceedings of the Aristotelian Society*, 59: 83–104. doi:10.1093/aristotelian/59.1.83
- , 1958b, “Moral Arguments”, *Mind*, 67(268): 502–513. doi:10.1093/mind/LXVII.268.502
- , 1972, “Morality as a System of Hypothetical Imperatives”, *The Philosophical Review*, 81(3): 305–316. doi:10.2307/2184328
- Frankena, William, 1963, “Recent Conceptions of Morality”, in G. Nakhnikian and H. Castañeda (eds.), *Morality and the Language of Conduct*, Detroit, MI: Wayne State University Press, pp. 1–24.
- , 1973, *Ethics*, Englewood Cliffs, N.J.: Prentice-Hall.
- , 1980, *Thinking about Morality*, Ann Arbor, MI: University of Michigan Press.
- Gert, Bernard, 2005, *Morality: Its Nature and Justification*, Revised Edition, New York: Oxford University Press.
- Gibbard, Allan, 1990, *Wise Choices, Apt Feelings*, Cambridge, MA: Harvard University Press.
- Goldman, Alan H., 2009, *Reasons from Within: Desires and Values*, New York: Oxford University Press. doi:10.1093/acprof:oso/9780199576906.001.0001
- Gray, Kurt, Liane Young, and Adam Waytz, 2012, “Mind Perception Is the Essence of Morality”, *Psychological Inquiry*, 23(2): 101–124. doi:10.1080/1047840X.2012.651387
- Greene, Joshua, 2013, *Moral Tribes: Emotion, Reason, and The Gap*


- between Us and Them*, New York: Penguin.
- Haidt, Jonathan, 2006, *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom*, New York: Basic Books.
- , 2011, *The Righteous Mind: Why Good People Are Divided by Politics and Religion*, New York: Pantheon.
- Haidt, Jonathan and Selin Kesebir, 2010, “Morality”, in S. Fiske, D. Gilbert, and G. Lindzey (eds.), *Handbook of Social Psychology*, 5th Edition, Hoboken, NJ: Wiley, pp. 797–832.
- Hare, R.M., 1952, *The Language of Morals*, New York: Oxford University Press.
- , 1963, *Freedom and Reason*, New York: Oxford University Press.
- , 1981, *Moral Thinking*, New York: Oxford University Press.
- Harman, Gilbert, 1975, “Moral Relativism Defended”, *The Philosophical Review*, 84(1): 3–22. doi:10.2307/2184078
- Hauser, Marc, 2006, *Moral Minds: How Nature Designed our Universal Sense of Right and Wrong*, New York: Harper Collins.
- Hobbes, Thomas, 1660 [1994], *Leviathan*, edited by Edwin Curly, Indianapolis: Hackett Publishing Company, 1994.
- , 1658 and 1651 [1991], *Man and Citizen*, (translations of six chapters of *De Homine* (1658) and all of *De Cive* (1651)), edited by Bernard Gert, Indianapolis: Hackett Publishing Company, 1991.
- Hooker, Brad, 2001, *Ideal Code, Real World: A Rule Consequentialist theory of Morality*, Oxford: Clarendon Press.
- Hume, David, 1751 [1975], *Enquiries concerning Human Understanding and concerning the Principles of Morals*, edited by L.A. Selby-Bigge, 3rd edition revised by P.H. Nidditch, Oxford: Clarendon Press, 1975.
- Kant, Immanuel, 1785 and 1797 [1993], *Groundwork of the Metaphysics of Morals: with On a Supposed Right to Lie because of Philanthropic Concerns*, 3rd edition, translated by J. Ellington, Indianapolis: Hackett, 1993.


- Kelly, Daniel, Stephen Stich, Kevin J. Haley, Serena J. Eng, and Daniel M. T. Fessler, 2007, “Harm, Affect, and the Moral/Conventional Distinction”, *Mind & Language*, 22(2): 117–131. doi:10.1111/j.1468-0017.2007.00302.x
- Klenk, Michael, 2019, “Moral Philosophy and the ‘Ethical Turn’ in Anthropology”, *Zeitschrift Für Ethik Und Moralphilosophie*, 2(2): 331–353. doi:10.1007/s42048-019-00040-9
- Laidlaw, James, 2016, “The Interactional Foundations of Ethics and the Formation and Limits of Morality Systems”, *HAU: Journal of Ethnographic Theory*, 6(1): 455–461. doi:10.14318/hau6.1.024
- Liao, S. Matthew (ed.), 2016, *Moral Brains. The Neuroscience of Morality*, New York: Oxford University Press.
doi:10.1093/acprof:oso/9780199357666.001.0001
- , 1999, *Dependent Rational Animals*, Chicago: Open Court.
- Machery, Edouard, 2012, “Delineating the Moral Domain”, *Baltic International Yearbook of Cognition, Logic and Communication*, 7(1): 1–14. doi:10.4148/biyclc.v7i0.1777
- Machery, Edouard and Ron Mallon, 2010, “The Evolution of Morality”, in Doris and The Moral Psychology Research Group 2010: 3–46.
- MacIntyre, Alasdair, 1957, “What Morality Is Not”, *Philosophy*, 32(123): 325–335. doi:10.1017/S0031819100051950
- , 1999, *Dependent Rational Animals*, Chicago: Open Court.
- Mikhail, John, 2007, “Universal Moral Grammar: Theory, Evidence and the Future”, *Trends in Cognitive Sciences*, 11(4): 143–152.
doi:10.1016/j.tics.2006.12.007
- Mill, John Stuart, 1861 [2002], *Utilitarianism*, edited by G. Sher, Indianapolis: Hackett, 2002.
- Moore, G.E., 1912, *Ethics*, New York: H. Holt.
- , 1903, *Principia Ethica*, New York: Cambridge University Press, 1993.
- Prinz, Jesse, 2007, *The Emotional Construction of Morals*, Oxford: Clarendon Press.
- Prinz, Jesse and Shaun Nichols, 2010, “Moral Emotions”, in Doris and The Moral Psychology Research Group 2010: 111–146.
- Rawls, John, 1971, *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Roedder, Erica and Gilbert Harman, 2010, “Linguistics and Moral Theory”, in Doris and The Moral Psychology Research Group 2010: 273–296.
- Scanlon, T. M., 1982, “Contractualism and Utilitarianism”, in *Utilitarianism and Beyond*, Amartya Sen and Bernard Williams (eds.), Cambridge: Cambridge University Press, 103–128.
doi:10.1017/CBO9780511611964.007
- , 1998, *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
- , 2011, “What is Morality?” in J. Shephard, S. Kosslyn, and E. Hammonds (eds.), *The Harvard Sampler: Liberal Education for the Twenty-First Century*, Cambridge, MA: Harvard University Press, pp. 243–66.
- Sidgwick, Henry, 1874, *Methods of Ethics*, Indianapolis: Hackett Pub. Co., 1981.
- Singer, Peter, 1993, *Practical Ethics*, 2nd Edition, Cambridge: Cambridge University Press.
- Sinnott-Armstrong, Walter (ed.), 2008, *Moral Psychology Volume 1, The Evolution of Morality: Adaptations and Innateness*, Cambridge, MA: MIT Press.
- , 2016, “The Disunity of Morality”, in Liao 2016: 331–354.
- Skorupski, John, 1993, “The Definition of Morality”, *Royal Institute of Philosophy Supplement*, 35: 121–144.
doi:10.1017/S1358246100006299
- Smart, J. J. C., 1956, “Extreme and Restricted Utilitarianism”, *The Philosophical Quarterly*, 6(25): 344–354. doi:10.2307/2216786

- Smith, Michael, 1994, *The Moral Problem*, Oxford: Blackwell.
- Sprigge, Timothy L. S., 1964, “Definition of a Moral Judgment”, *Philosophy*, 39(150): 301–322. doi:10.1017/S0031819100055777
- Strawson, P. F., 1961, “Social Morality and Individual Ideal”, *Philosophy*, 36(136): 1–17. doi:10.1017/S003181910005779X
- Thomson, J.J. and G. Dworkin (eds.), 1968, *Ethics*, New York: Harper & Row.
- Toulmin, Stephen, 1950, *An Examination of the Place of Reason in Ethics*, Cambridge: Cambridge University Press.
- Turiel, Elliot, 1983, *The Development of Social Knowledge: Morality and Convention*, Cambridge: Cambridge University Press.
- Warnock, Geoffrey, 1971, *The Object of Morality*, London: Methuen.
- Westermarck, Edward, 1960, *Ethical Relativity*, Paterson, N.J.: Littlefield, Adams.
- Williams, Bernard, 1985, *Ethics and the Limits of Philosophy*, London: Fontana.
- Wong, David B., 1984, *Moral Relativity*, Berkeley CA: University of California Press.
- , 2006, *Natural Moralities: A Defense of Pluralistic Relativism*, New York: Oxford University Press. doi:10.1093/0195305396.001.0001
- , 2014, “Integrating Philosophy with Anthropology in an Approach to Morality”, *Anthropological Theory*, 14(3): 336–355. doi:10.1177/1463499614534554
- Wren, T.E. (ed.), 1990, *The Moral Domain: Essays in the Ongoing Discussion Between Philosophy and the Social Sciences*, Cambridge, MA: MIT Press.

Academic Tools

-  How to cite this entry.
-  Preview the PDF version of this entry at the Friends of the SEP Society.

 Look up this entry topic at the Internet Philosophy Ontology Project (InPhO).

 Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

[Please contact the author with suggestions.]

Related Entries

consequentialism | ethics: natural law tradition | Hobbes, Thomas: moral and political philosophy | Kant, Immanuel | Mill, John Stuart | moral realism | moral relativism | moral skepticism

Copyright © 2020 by the authors
Bernard Gert and Joshua Gert