

PDF version of the entry  
Moral Theory  
<https://plato.stanford.edu/archives/fall2022/entries/moral-theory/>  
from the FALL 2022 EDITION of the

## STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Co-Principal Editors: Edward N. Zalta & Uri Nodelman  
Associate Editors: Colin Allen, Hannah Kim, & Paul Oppenheimer  
Faculty Sponsors: R. Lanier Anderson & Thomas Icard  
Editorial Board: <https://plato.stanford.edu/board.html>  
Library of Congress 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 © 2022 by the publisher  
The Metaphysics Research Lab  
Department of Philosophy  
Stanford University, Stanford, CA 94305

Moral Theory  
Copyright © 2022 by the author  
Julia Driver

All rights reserved.

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

## Moral Theory

*First published Mon Jun 27, 2022*

There is much disagreement about what, exactly, constitutes a moral theory. Some of that disagreement centers on the issue of demarcating the *moral* from other areas of practical normativity, such as the ethical and the aesthetic. Some disagreement centers on the issue of what a moral *theory's* aims and functions are. In this entry, both questions will be addressed. However, this entry is about moral theories as *theories*, and is not a survey of specific theories, though specific theories will be used as examples.

- 1. Morality
  - 1.1 Common-sense Morality
  - 1.2 Contrasts Between Morality and Other Normative Domains
- 2. Theory and Theoretical Virtues
  - 2.1 The Tasks of Moral Theory
  - 2.2 Theory Construction
- 3. Criteria
- 4. Decision Procedures and Practical Deliberation
- Bibliography
- Academic Tools
- Other Internet Resources
- Related Entries

---

## 1. Morality

When philosophers engage in moral theorizing, what is it that they are doing? Very broadly, they are attempting to provide a systematic account of morality. Thus, the object of moral theorizing is morality, and, further, morality as a normative system.

At the most minimal, morality is a set of norms and principles that govern our actions with respect to each other and which are taken to have a special kind of weight or authority (Strawson 1961). More fundamentally, we can also think of morality as consisting of moral reasons, either grounded in some more basic value, or, the other way around, grounding value (Raz 1999).

It is common, also, to hold that moral norms are *universal* in the sense that they apply to and bind everyone in similar circumstances. The principles expressing these norms are also thought to be *general*, rather than specific, in that they are formulable “without the use of what would be intuitively recognized as proper names, or rigged definite descriptions” (Rawls 1979, 131). They are also commonly held to be *impartial*, in holding everyone to count equally.

## 1.1 Common-sense Morality

... Common-sense is... an exercise of the judgment unaided by any Art or system of rules : such an exercise as we must necessarily employ in numberless cases of daily occurrence ; in which, having no established principles to guide us ... we must needs act on the best extemporaneous conjectures we can form. He who is eminently skillful in doing this, is said to possess a superior degree of Common-Sense. (Richard Whatley, *Elements of Logic*, 1851, xi–xii)

“Common-Sense Morality”, as the term is used here, refers to our *pre-theoretic* set of moral judgments or intuitions or principles.<sup>[1]</sup> When we engage in theory construction (see below) it is these common-sense intuitions that provide a touchstone to theory evaluation. Henry Sidgwick believed that the principles of Common-Sense Morality were important in helping us understand the “first” principle or principles of morality.<sup>[2]</sup>

Indeed, some theory construction explicitly appeals to puzzles in common-sense morality that need resolution – and hence, need to be addressed theoretically.

Features of commons sense morality are determined by our normal reactions to cases which in turn suggest certain normative principles or insights. For example, one feature of common-sense morality that is often remarked upon is the self/other asymmetry in morality, which manifests itself in a variety of ways in our intuitive reactions. For example, many intuitively differentiate morality from prudence in holding that morality concerns our interactions with others, whereas prudence is concerned with the well-being of the individual, from that individual’s point of view.

Also, according to our common-sense intuitions we are allowed to pursue our own important projects even if such pursuit is not “optimific” from the impartial point of view (Slote 1985). It is also considered permissible, and even admirable, for an agent to sacrifice her own good for the sake of another even though that is not optimific. However, it is impermissible, and outrageous, for an agent to similarly sacrifice the well-being of another under the same circumstances. Samuel Scheffler argued for a view in which consequentialism is altered to include agent-centered prerogatives, that is, prerogatives to not act so as to maximize the good (Scheffler 1982).

Our reactions to certain cases also seem to indicate a common-sense commitment to the moral significance of the distinction between intention and foresight, doing versus allowing, as well as the view that distance between agent and patient is morally relevant (Kamm 2007).

Philosophers writing in empirical moral psychology have been working to identify other features of common-sense morality, such as how prior moral evaluations influence how we attribute moral responsibility for actions (Alicke et. al. 2011; Knobe 2003).

What many ethicists agree upon is that common-sense is a bit of a mess. It is fairly easy to set up inconsistencies and tensions between common-sense commitments. The famous Trolley Problem thought experiments illustrate how situations which are structurally similar can elicit very different intuitions about what the morally right course of action would be (Foot 1975). We intuitively believe that it is worse to kill someone than to simply let the person die. And, indeed, we believe it is wrong to kill one person to save five others in the following scenario:

David is a great transplant surgeon. Five of his patients need new parts—one needs a heart, the others need, respectively, liver, stomach, spleen, and spinal cord—but all are of the same, relatively rare, blood-type. By chance, David learns of a healthy specimen with that very blood-type. David can take the healthy specimen's parts, killing him, and install them in his patients, saving them. Or he can refrain from taking the healthy specimen's parts, letting his patients die. (Thomson 1976, 206)

And yet, in the following scenario we intuitively view it entirely permissible, and possibly even obligatory, to kill one to save five:

Edward is the driver of a trolley, whose brakes have just failed. On the track ahead of him are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Edward can turn the trolley onto it. Unfortunately there is one person on the right-hand track. Edward can turn the trolley, killing the one; or he can refrain from turning the trolley, killing the five. (Thomson 1976, 206).

*Theorizing* is supposed to help resolve those tensions in a principled way. *Theory construction* attempts to provide guidance in how to resolve such tensions and how to understand them.

## 1.2 Contrasts Between Morality and Other Normative Domains

### 1.2.1 Morality and Ethics

Ethics is generally understood to be the study of “living well as a human being”. This is the topic of works such as Aristotle’s *Nicomachean Ethics*, in which the aim of human beings is to exemplify human excellence of character. The sense in which we understand it here is that ethics is broader than morality, and includes considerations of personal development of oneself and loved ones. This personal development is important to a life well lived, intuitively, since our very identities are centered on projects that we find important. Bernard Williams and others refer to these projects as “ground projects”. These are the sources of many of our reasons for acting. For Williams, if an agent seeks to adopt *moral* considerations, or be guided by them, then important *ethical* considerations are neglected, such as personal integrity and authenticity (Williams 1977; Wolf 1982). However, Williams has a very narrow view of what he famously termed “the morality system” (Williams 1985).

Williams lists a variety of objectionable features of the morality system, including the *inescapability* of moral obligations, the *overridingness* of moral obligation, *impartiality*, and the fact that in the morality system there is a push towards *generalization*.

There has been considerable discussion of each of these features of the morality system, and since Williams, a great deal of work on the part of standard moral theorists on how each theory addresses the considerations he raised. Williams’ critique of the morality system was part of a general criticism of moral theory in the 1980s on the grounds of its uselessness, harmfulness, and even its impossibility (Clarke 1987). This anti-theory

trend was prompted by the same dissatisfaction with consequentialism and deontology that led to the resurgence of Virtue Ethics.

A major criticism of this view is that it has a very narrow view of what counts as a moral theory. Thus, some of these approaches simply rejected some features of Williams's characterization of the morality system, such as impartiality. Others, however, Williams' included, attacked the very project of moral theory. This is the 'anti-theory' attack on moral theorizing. For example, Annette Baier argued that morality cannot be captured in a system of rules, and this was a very popular theme amongst early virtue ethicists. On this view, moral theory which systematizes and states the moral principles that ought to guide actions is simply impossible: "Norms in the form of virtues may be essentially imprecise in some crucial ways, may be mutually referential, but not hierarchically orderable, may be essentially self-referential" (Baier 220).

Robert Louden even argued that the best construal of virtue ethics is not as an ethical theory, but as anti-theory that should not be evaluated as attempting to theorize morality at all. (Louden 1990). According to Louden, moral theories are formulated to a variety of reasons, including to provide solutions to problems, formulas for action, universal principles, etc. Louden notes that this characterization is very narrow and many would object to it, but he views anti-theory not so much as a position against *any* kind of moral theorizing, but simply the kind that he viewed as predominant prior to the advent of Virtue Ethics. This is a much less severe version of anti-theory as it, for example, doesn't seem to regard weightiness or importance of moral reasons as a problem.

Some of the problems that Williams and other anti-theorists have posed for morality, based on the above characteristics, are:

1. Morality is too demanding and pervasive: that is, the view that moral reasons are weighty indicates that we should be giving them priority

over other sorts of reasons. Further, they leach into all aspects of our lives, leaving very little morally neutral.

2. Morality is alienating. There are a variety of ways in which morality can be alienating. As Adrian Piper notes, morality might alienate the agent from herself or might alienate the agent from others – impartiality and universality might lead to this, for example (Piper 1987; Stocker 1976). Another way we can understand alienation is that the agent is alienated from the true justifications of her own actions – this is one way to hold that theories which opt for indirection can lead to alienation (see section 4 below).
3. Morality, because it is impartial, makes no room for special obligations. That is, if the right action is the one that is impartial between persons, then it does not favor the near and dear. On this picture it is difficult to account for the moral requirements that parents have towards their own children, and friends have towards each other. These requirements are, by their nature, not impartial.
4. Morality is committed to providing guides for action that can be captured in a set of rules or general principles. That is, morality is codifiable and the rules of morality are general.

Morality requires too much. The basic worry is that the morality system is voracious and is creeping into all aspects of our lives, to the detriment of other important values. The worry expressed by 4 takes a variety of forms. For example, some take issue with a presupposition of 4, arguing that there are no moral principles at all if we think of these principles as *guiding action*. Some argue that there are no moral principles that are complete, because morality is not something that is *codifiable*. And, even if morality was codifiable, the 'principles' would be extremely *specific*, and not qualify as principles at all.

Since Williams' work, philosophers have tried to respond to the alienation worry by, for example, providing accounts of the ways in which a person's reasons can guide without forming an explicit part of practical deliberation. Peter Railton, for example, argues in favor of a form of objective consequentialism, *Sophisticated Consequentialism*, in which the rightness of an action is a function of its actual consequences (Railton 1984). On Railton's view, one can be a good consequentialist without being alienated from loved ones. Though not attempting to defend moral theory *per se*, other writers have also provided accounts of how agents can act on the basis of reasons – and thus perform morally worthy actions, even though these reasons are not explicitly articulated in their practical deliberations (Arpaly 2002; Markovits 2014). Deontologists have argued that autonomous action needn't involve explicit invocation of, for example, the Categorical Imperative (Herman 1985). Generally, what characterizes these moves is the idea that the justifying reasons are present in some form in the agent's psychology – they are recoverable from the agent's psychology – but need not be explicitly articulated or invoked by the agent in acting rightly.

One way to elaborate on this strategy is to argue that the morally good agent is one who responds to the right sorts of reasons, even though the agent can't articulate the nature of the response (Arpaly 2002). This strategy makes no appeal to codifiable principles, and is compatible with a wide variety of approaches to developing a moral theory. It relies heavily on the concept, of course, of "reason" and "moral reason," which many writers on moral issues take to be fundamental or basic in any case.

There has also been debate concerning the proper scope of morality, and how moral theories can address problems relating to impartiality. Kant and the classical utilitarians believed that moral reasons are impartial, what others have termed agent-neutral. Indeed, this is one point of criticism that virtue ethics has made of these two theories. One might argue that moral

reasons are impartial, but that there are other reasons that successfully compete with them – reasons relating to the near and dear, for example, or one's own ground projects. Or, one could hold that morality includes special reasons, arising from special obligations, that also morally justify our actions.

The first strategy has been pursued by Bernard Williams and other "anti-theorists". Again, Williams argues that morality is a special system that we would be better off without (Williams 1985). In the morality system we see a special sense of "obligation" – moral obligation – which possesses certain features. For example, moral obligation is *inescapable* according to the morality system. A theory such as Kant's, for example, holds that we *must* act in accordance with the Categorical Imperative. It is not optional. This is because morality is represented as having authority over us in ways that even demand sacrifice of our personal projects, of the very things that make our lives go well for us. This seems especially clear for Utilitarianism, which holds that we must *maximize* the good, and falling short of maximization is *wrong*. A Kantian will try to avoid this problem by appealing to obligations that are less demanding, the imperfect ones. But, as Williams points out, these are still *obligations*, and as such can only be overridden by other obligations. Thus, the theories also tend to present morality as *pervasive* in that morality creeps into every aspect of our lives, making no room for neutral decisions. For example, even decisions about what shoes to wear to work becomes a moral one:

Once the journey into more general obligations has started, we may begin to get into trouble – not just philosophical trouble, but the conscience trouble – with finding room for morally indifferent actions. I have already mentioned the possible moral conclusion that one may take some particular course of action. That means that there is nothing else I am obliged to do. But if we have accepted general and indeterminate obligations to further various

moral objectives...they will be waiting to provide work for idle hands... (Williams 1985, 181)

He goes on to write that in order to get out of this problem, "...I shall need one of those fraudulent items, a duty to myself" (Williams 1985, 182). Kantian Ethics does supply this. Many find this counterintuitive, since the self/other asymmetry seems to capture the prudence/morality distinction, but Kantians such as Tom Hill, jr. have made strong cases for at least *some* moral duties to the self. In any case, for writers such as Williams, so much the worse for *morality*.

Other writers, also concerned about the problems that Williams has raised argue, instead, that morality does make room for our partial concerns and projects, such as the norms governing our relationships, and our meaningful projects. Virtue ethicists, for example, are often comfortable pointing out that morality is not thoroughly impartial because there are virtues of partiality. Being a good mother involves having a preference for the well-being of one's own children. The mother who really is impartial would be a very bad mother, lacking in the appropriate virtues.

Another option is to hold that there are partial norms, but those partial norms are themselves justified on impartial grounds. This can be spelled out in a variety of different ways. Consider Marcia Baron's defense of impartiality, where she notes that critics of impartiality are mistaken because they confuse *levels* of justification: "Critics suppose that impartialists insisting on impartiality at the level of rules or principles are committed to insisting on impartiality at the level of deciding what to do in one's day-to-day activities" (Baron 1991). This is a mistake because impartialists can justify partial norms by appealing to impartial rules or principles. She is correct about this. Even Jeremy Bentham believed, for example, that the principle of utility ought not be applied in every case, though he mainly appealed to efficiency costs of using the principle all the

time. But one can appeal to other considerations. Frank Jackson uses an analogy with predators to argue that partial norms are strategies for maximizing the good, they offer the best chance of actually doing so given our limitations (Jackson 1991). Similarly, a Kantian such as Tom Hill, jr., as Baron notes, can argue that impartiality is part of an ideal, and ought not govern our day-to-day lives (Hill 1987). Does this alienate people from others? The typical mother shows the right amount of preference for her child, let's say, but doesn't herself think that this is justified on the basis of promoting the good, for example. A friend visits another in the hospital and also does not view the partiality as justified by any further principles. But this is no more alienating than someone being able to make good arguments and criticize bad ones without a knowledge of inference rules. Maybe it is better to have an awareness of the underlying justification, but for some theories even that is debatable. For an objective theorist (see below) it may be that knowing the underlying justification can interfere with doing the right thing, in which case it is better not to know. For some theorists, however, such as neo-Aristotelian virtue ethicists, a person is not truly virtuous without such knowledge and understanding, though Rosalind Hursthouse (1999) does *not* make this a requirement of right action.

Recently consequentialists have been approaching this issue through the theory of value itself, arguing that there are agent-relative forms of value. This approach is able to explain the intuitions that support partial moral norms while retaining the general structure of consequentialism (Sen 2000). Douglas Portmore, for example, argues for a form of consequentialism that he terms "commonsense consequentialism" as it is able to accommodate many of our everyday moral intuitions (Portmore 2011). He does so by arguing that (1) the deontic status of an act, whether it is right or wrong, is determined by what reasons the agent has for performing it – if an agent has a decisive reason to perform the act in question, then it is morally required. Combined with (2) a teleological

view of practical reasons in which our reasons for performing an action are a function of what we have reason to prefer or desire we are led to a form of act-consequentialism but one which is open to accepting that we have reason to prefer or desire the well-being of the near and dear over others.

Though much of this is controversial, there is general agreement that moral reasons are *weighty*, are *not* egoistic – that is, to be contrasted with prudential reasons, and are concerned with issues of value [duty, fittingness].

### 1.2.2. Morality and Aesthetics

Moral modes of evaluation are distinct from the aesthetic in terms of their content, but also in terms of their authority. So, for example, works of art are evaluated as “beautiful” or “ugly”, and those evaluations are not generally considered as universal or as objective as moral evaluations. These distinctions between moral evaluation and aesthetic evaluation have been challenged, and are the subject of some interesting debates in metaethics on the nature of both moral and aesthetic norms and the truth-conditions of moral and aesthetic claims. But, considered intuitively, aesthetics seems at least *less* objective than morality.

A number of writers have noted that we need to be cognizant of the distinction between moral norms and the norms specific to other normative areas in order to avoid fallacies of evaluation, and much discussion has centered on a problem in aesthetics termed the “Moralistic Fallacy” (D’Arms and Jacobson 2000).

One challenge that the anti-theorists have raised for morality was to note that in a person’s life there will be certain norm clashes – including clashes between *types* of norms such as the moral and the aesthetic. It is

giving too much prominence to the moral that judges a person’s life as going well relative to the fulfillment or respect of *those* norms. Can’t a human life go well, even when that life sacrifices morality for aesthetics?

This sort of debate has a long history in moral theory. For example, it arose as a form of criticism of G. E. Moore’s Ideal Utilitarianism, which treated beauty as an intrinsic good, and rendering trade-offs between behaving well towards others and creating beauty at least in principle justified *morally* (Moore 1903). But the anti-theorists do not pursue this method of accommodating the aesthetic, instead arguing that it is a separate normative realm which has its own weight and significance in human flourishing.

## 2. Theory and Theoretical Virtues

There is agreement that theories play some kind of *systematizing* role, and that one function is to examine important concepts relevant to morality and moral practice and the connections, if any, between them. For example, one very common view in the middle of the 20<sup>th</sup> century, attributed to John Rawls, was to view moral theory as primarily interested in understanding the ‘right’ and the ‘good’ and connections between the two (Rawls). Priority claims are often a central feature in the systematizing role of moral theory. Related to this is the issue of explanatory, or theoretical, *depth*. That is, the deeper the explanation goes, the better.

Theories also strive for *simplicity*, *coherence*, and *accuracy*. The fewer epicycles the theory has to postulate the better, the parts of the theory should fit well together. For example, the theory should not contain inconsistent principles, or have inconsistent implications. The theory should cover the phenomena in question. In the case of moral theories, the phenomena in question are thought to be our considered moral intuitions

or judgements. Another coherence condition involves the theory cohering with a person's set of considered judgments, as well.

One last feature that needs stressing, particularly for moral theories, is *applicability*. One criticism of some normative ethical theories is that they are not applicable. For example, Virtue Ethics has been criticized for not providing an account of what our moral *obligations* are – appealing to what the virtuous person would do in the circumstances would seem to set a very high bar or doesn't answer the relevant question about how we should structure laws guiding people on what their social obligations are. Similarly, objective consequentialists, who understand "right action" in terms of actual consequences have been criticized for rendering what counts as a right action in a given circumstance unknowable, and thus useless as a guide to action. Both approaches provide responses to this worry, but this supports the claim that a desideratum of a moral theory is that it be applicable.

## 2.1 The Tasks of Moral Theory

One task (though this is somewhat controversial) of a moral theory is to give an account of right actions. Often, this will involve an explication of what counts as good – some theories then get spelled out in terms of how they approach the good, by maximizing it, producing enough of it, honoring it, etc. In addition, some theories explicate the right in terms of acting in accordance with one's duties, or acting as a virtuous person would act. In these cases the notions of 'duty' and 'virtue' become important to the overall analysis, and one function of moral theory is to explore the systematic connections between duty or virtue and the right and the good.

Moral theories also have both substantive and formal aims. Moral theories try to provide criteria for judging actions. It might be that the criterion is

simple, such as right actions maximize the good, or it may be complex, such as the right action is the one that gives adequate weight to each competing duty. Sometimes, in recognition that there is not always "the" right action, the theory simply provides an account of wrongness, or permissibility and impermissibility, which allows that a range of actions might count as "right".

In addition to simply providing criteria for right or virtuous action, or for being a virtuous person, a given moral theory, for example, will attempt to explain *why* something, like an action or character trait, has a particular moral quality, such as rightness or virtuousness. Some theories view rightness as grounded in or explained by *value*. Some view rightness as a matter of *reasons* that are prior to value. In each case, to provide an explanation of the property of 'rightness' or 'virtuousness' will be to provide an account of what the grounding value is, or an account of reasons for action.

In addition, moral theories may also provide *decision-procedures* to employ in determining how to act rightly or virtuously, conditions on being good or virtuous, or conditions on morally appropriate practical deliberation. Thus, the theory provides substance to evaluation and reasons. However, moral theories, in virtue of providing an explanatory framework, help us see connections between criteria and decision-procedures, as well as provide other forms of systemization. Thus, moral theories will be themselves evaluated according to their theoretical virtues: simplicity, explanatory power, elegance, etc. To evaluate moral theories *as theories*, each needs to be evaluated in terms of how well it succeeds in achieving these theoretical goals.

There are many more specialized elements to moral theories as well. For example, a moral theory often concerns itself with features of moral psychology relevant to action and character, such as motives, intentions,

emotions, and reasons responsiveness. A moral theory that incorporates consideration of consequences into the determination of moral quality, will also be concerned with issues surrounding the proper *aggregation* of those consequences, and the *scope* of the consequences to be considered.

## 2.2 Theory Construction

There's been a long history of comparing moral theories to other sorts of theories, such as scientific ones. For example, in meta-ethics one issue has to do with the nature of moral "evidence" on analogy with scientific evidence. On what Ronald Dworkin terms the "natural model" the truths of morality are discovered, just as the truths of science are (Dworkin 1977, 160). It is our considered intuitions that provide the clues to discover these moral truths, just as what is observable to us provides the evidence to discover scientific truths. He compared this model with the "constructive model" in which the intuitions themselves are features of the theory being constructed and are not analogous to observations of the external world.

Yet, even if we decide that morality lacks the same type of phenomena to be accounted for as science, morality clearly figures into our normative judgments and reactions. One might view these – our intuitions about moral cases, for example – to provide the basic data that needs to be accounted for by a theory on either model.

One way to "account for" our considered intuitions would be to debunk them. There is a long tradition of this in moral philosophy as well. When scholars provided *genealogies* of morality that explained our considered intuitions in terms of social or evolutionary forces that are not sensitive to the truth, for example, they were debunking morality by undercutting the authority of our intuitions to provide insight into it (Nietzsche 1887 [1998], Joyce 2001, Street 2006). In this entry, however, we consider the ways in which moral theorists have constructed their accounts by taking

the intuitions seriously as something to be systematized, explained, and as something that can be applied to generate the correct moral decisions or outcomes.

Along these lines, one method used in theory construction would involve the use of reflective equilibrium and inference to the best explanation. For example, one might notice an apparent inconsistency in moral judgements regarding two structurally similar cases and then try to figure out what principle or set of principles would achieve consistency between them. In this case, the theorist is trying to figure out what best explains both of those intuitions. But one also might, after thinking about principles one already accepts, or finds plausible, reject one of those intuitions on the basis of it not cohering with the rest of one's considered views. But full theory construction will go beyond this because of the fully theoretical virtues discussed earlier. We want a systematic account that coheres well not only with itself, but with other things that we believe on the basis of good evidence.

## 3. Criteria

Consider the following:

Malory has promised to take Chris grocery shopping. Unfortunately, as Malory is leaving the apartment, Sam calls with an urgent request: please come over to my house right now, my pipes have broken and I need help! Torn, Malory decides to help Sam, and thus breaks a promise to Chris.

Has Malory done the right thing? The virtuous thing? Malory has broken a promise, which is *pro tanto* wrong, but Sam is in an emergency and needs help right away. Even if it is clear that what Malory did was right in the circumstances, it is an interesting question as to why it is right. What can

we appeal to in making these sorts of judgments? This brings to light the issue of how one *morally* justifies one's actions. This is the task of understanding what the justifying reasons are for our actions. What makes an action *the thing to do* in the circumstances? This is the criterion of rightness (or wrongness). We will focus on the criterion of rightness, though the criterion issue comes up with other modes of moral evaluation, such as judging an action to be virtuous, or judging it to be good in some respect, even if not right. Indeed, some writers have argued that 'morally right' should be jettisoned from modern secular ethics, as it presupposes a conceptual framework left over from religiously based accounts which assume there is a God (Anscombe 1958). We will leave these worries aside for now, however, and focus on standard accounts of criteria.

The following are some toy examples that exhibit differing structural features for moral theories and set out different criteria:

*Consequentialism.*

The right action is the action that *produces* good amongst the options open to the agent at the time of action (Singer). The most well-known version of this theory is Classical Utilitarianism, which holds that the right action promotes pleasure (Mill).

*Kantian Deontology.*

The morally worthy action is in accordance with the Categorical Imperative, which requires an agent refrain from acting in a way that fails to respect the rational nature of other persons (Kant).

*Rossian Deontology.*

The right action is the action that best accords with the fulfillment and/or non-violation of one's *prima facie* duties (Ross).

*Contractualism.*

An action is morally wrong if it is an act that would be forbidden

by principles that rational persons could not reasonably reject (Scanlon).

*Virtue Ethics.*

The right action is the action that a virtuous person would characteristically perform in the circumstances (Hursthouse 1999).

These principles set out the *criterion* or *standard* for evaluation of actions. They do not necessarily tell us *how* to perform right actions, and are not, in themselves, decision-procedures, though they can easily be turned into decision procedures, such as: you ought to try to perform the action that maximizes the good amongst the options available to you at the time of action. This might not be, and in ordinary circumstance probably isn't, a very good decision-procedure, and would itself need to be evaluated according to the criterion set out by the theory.

These theories can be divided, roughly, into the deontological, consequentialist, and virtue ethical categories. There has been a lively debate about how, exactly, to delineate these categories. Some have held that deontological theories were just those theories that were not consequentialist. A popular conception of consequentialist theories is that they are reductionist in a particular way – that is, in virtue of reducing deontic features of actions (e.g. rightness, obligatoriness) to facts about an agent's options and the consequences of those options (Smith 2009). If that is the case, then it seems that deontological approaches are just the ones that are not reductive in this manner. However, this fails to capture the distinctive features of many forms of virtue ethics, which are neither consequentialist nor necessarily concerned with what we *ought to do*, our *duties* as opposed to what sorts of persons we should be.

One way to distinguish consequentialist from deontological theories is in terms of how each approaches value. Philip Pettit has suggested that while consequentialist theories required *promotion* of value, deontological

theories recommend that value be *honored* or *respected*. On each of these views, value is an important component of the theory, and theories will be partially delineated according to their theory of value. A utilitarian such as Jeremy Bentham believes that hedonism is the correct theory of value, whereas someone such as G. E. Moore, a utilitarian but a pluralist regarding value, believes that hedonism is much too narrow an account. A Kantian, on the other hand, views value as grounded in rational nature, in a will conforming to the Categorical Imperative.

Because of the systematizing function of moral theory discussed earlier, the simplest account is to be preferred and thus there is a move away from endorsing value pluralism. Of course, as intuitive pressure is put on each of the simpler alternatives, a pluralistic account of criteria for rightness and wrongness has the advantage of according best with moral intuitions.

Reasons-first philosophers will delineate the theories somewhat differently. For example, one might understand goodness as a matter of what we have reason to desire, in which case what we have reason to desire is prior to goodness rather than the other way around. Value is still an important component of the theories, it is simply that the value is grounded in reasons.

Another distinction between normative theories is that between subjective and objective versions of a type of theory. This distinction cuts across other categories. For example, there are subjective forms of all the major moral theories, and objective versions of many. An objective standard of right holds that the agent must actually meet the standard – and meeting the standard is something ‘objective’, not dependent on the agent’s psychological states – in order to count as right or virtuous. Subjective standards come in two broad forms:

1. *Psychology sensitive*: are the justifying reasons part of the agent’s deliberative processes? Or, more weakly, are they “recoverable” from

the agent’s psychology [perhaps, for example, the agent has a commitment to the values that provide the reasons].

2. *Evidence sensitive*: the right action isn’t the one that actually meets the standard, but instead, is the action that the agent could foresee would meet that standard. [there are many different ways to spell this out, depending on the degree of evidence that is relevant: in terms of what the agent actually foresees, what is foreseeable by the agent given what the agent knows, is foreseeable by someone in possession of a reasonable amount of evidence, etc.]

Of course, these two can overlap. For theorists who are *evaluational internalists*, evidence-sensitivity doesn’t seem like a plausible way of spelling out the standard, except, perhaps, indirectly. The distinction frequently comes up in Consequentialism, where the Objective standard is taken to be something like: the right action is the action that actually promotes the good and the Subjective standard is something like: the right action is the action that promotes the good by the agent’s own lights (psychology sensitive) or the right action is the action that promotes the foreseeable good, given evidence available at the time of action (evidence sensitive standard). It is certainly possible for other moral standards to be objective. For example, the right action is the action that the virtuous person would perform, even though the agent does not realize it is what the virtuous agent would do in the circumstances, and even if the person with the best available evidence couldn’t realize it is what the virtuous person would do in the circumstances.

We certainly utter locutions that support both subjective and objective uses of what we ‘ought’ to do, or what is ‘right’. Frank Jackson notes this when he writes:

- ...we have no alternative but to recognize a whole range of oughts
- what she ought to do by the light of her beliefs at the time of

action, ...what she ought to do by the lights of one or another onlooker who has different information on the subject, and, what is more, what she ought to do by God's lights...that is, by the lights of one who *knows* what will and would happen for each and every course of action. (Jackson 1991, 471).

For Jackson, the primary ought, the primary sense of 'rightness' for an action, is the one that is "most immediately relevant to action" since, otherwise, we have a problem of understanding how the action is the agent's. Thus, the subjective 'ought' is primary in the sense that this is the one that ethical theory should be concerned with (Jackson 1991). Each type of theorist makes use of our ordinary language intuitions to make their case. But one desideratum of a theory is that it not simply reflect those intuitions, but also provides the tools to critically analyze them. Given that our language allows for both sorts of 'ought,' the interesting issue becomes which, if either, has primacy in terms of actually providing the standard by which other things are evaluated? Moral theory needn't only be concerned with what the right action is from the agent's point of view.

There are three possibilities:

1. neither has primacy
2. the subjective has primacy
3. the objective has primacy

First off we need to understand what we mean by "primacy". Again, for Frank Jackson, the primary sense of 'right' or 'ought' is subjective, since what we care about is the 'right' that refers to an inward story, the story of our agency, so to speak. On this view, the objective and subjective senses may have no relationship to each other at all, and which counts as primary simply depends upon our interests. However, the issue that concerns us here is whether or not one sense can be *accounted for* in terms of the other.

Option 1 holds that there is no explanatory connection. That is not as theoretically satisfying. Option 2 holds either there really is no meaningful objective sense, just the subjective sense, or the objective sense is understood in terms of the subjective.

Let's look at the objective locution again "He did the right thing, but he didn't know it at the time (or he had no way of knowing it at the time)". Perhaps all this means is "He did what someone with all the facts and correct set of values would have judged right by their own lights" – this would be extensionally the same as "He performed the action with the best actual consequences". This is certainly a possible account of what objective right means which makes use of a subjective standard. But it violates the spirit of the subjective standard, since it ties rightness neither to the psychology of the agent, or the evidence that is actually available to the agent. For that reason, it seems more natural to opt for 3. An advantage of this option is that gives us a nice, unified account regarding the connection between the objective and the subjective. Subjective standards, then, are standards of praise and blame, which are themselves evaluable according to the objective standard. Over time, people are in a position to tell whether or not a standard actually works in a given type of context. Or, perhaps it turns out that there are several standards of blame that differ in terms of severity. For example, if someone acts negligently a sensible case can be made that the person is blameworthy but not as blameworthy as if they had acted intentionally.

As to the worry that the objective standard doesn't provide action guidance, the objective theorist can hold that action guidance is provided by the subjective standards of praise/blameworthiness. Further, the standard itself can provide what we need for action guidance through normative review (Driver 2012). Normative review is a retrospective look at what does in fact meet the standard, and under what circumstances.

Now, consider a virtue ethical example. The right action is the action that is the actual action that a virtuous person would perform characteristically, in the circumstances, rather than the action that the agent believes is the one the virtuous person would perform. Then we evaluate an agent's "v-rules" in terms of how close they meet the virtuous ideal.

#### 4. Decision Procedures and Practical Deliberation

Another function of moral theory is to provide a decision procedure for people to follow so as to best insure they perform right actions. Indeed, some writers, such as R. M. Hare hold action guidance to be the function of the moral principles of the theory (Hare 1965). This raises the question of what considerations are relevant to the content of such principles – for example, should the principles be formulated taking into account the epistemic limitations of most human beings? The requirement that moral principles be action guiding is what Holly Smith terms the "Useability Demand": "...an acceptable moral principle must be useable for guiding moral decisions..." (Smith 2020, 11). Smith enumerates different forms satisfaction of this demand can take, and notes that how one spells out a principle in order to meet the demand will depend upon how the moral theorist views moral success. For example, whether or not success is achieved in virtue of simply making the right decision or if, in addition to making the right decision, the agent must also have successful follow-through on that decision.

There has been enormous debate on the issue of what is involved in following a rule or principle, and some skepticism that this is in fact what we are doing when we *take ourselves* to be following a rule. (Kripke 1982) Some virtue theorists believe that it is moral perception that actually does the guiding, and that a virtuous person is able to perceive what is morally relevant and act accordingly (McDowell 1979).

As discussed earlier in the section on criteria, however, this is also controversial in that some theorists believe that decision procedures themselves are not of fundamental significance. Again, objective consequentialist who believes that the fundamental task of theory is to establish a criterion for right argues that decision procedures will themselves be established and evaluated on the basis of how well they get us to actually achieving the right. Thus, the decision-procedures are derivative. Others, such as subjective consequentialists, will argue that the decision-procedures *specify* the criterion in the sense that following the decision-procedure itself is sufficient for meeting the criterion. For example, an objective consequentialist will hold that the right action maximizes the good, whereas the subjective consequentialist might hold that the right action is to *try* to maximize the good, whether or not one actually achieves it (Mason 2003 and 2019). *Following* the decision-procedure itself, then, is the criterion.

The distinction between criterion and decision-procedure has been acknowledged and discussed at least since Sidgwick, though it was also mentioned by earlier ethicists. This distinction allows ethical theories to avoid wildly implausible implications. For example, if the standard that the theory recommends is 'promote the good' it would be a mistake to think that 'promote the good' *needs* to be part of the agent's deliberation. The consequentialist might say that, instead, it is an empirical issue as to what the theory is going to recommend as a decision-procedure, and that recommendation could vary from context to context. There will surely be circumstances in which it would be best to think in terms of meeting the standard itself, but again that is an empirical issue. Likewise, it is open to a Virtue Ethicist to hold that the right action is the one the virtuous agent would perform in the circumstances, but also hold that the agent's deliberative processes need not make reference to the standard. Pretty much all theories will want to make some space between the standard and the decision-procedure in order to avoid a requirement that agent's must

think in terms of the correct standard, in order to act rightly, or even act with moral worth. There is a distinction to be made between doing the right thing, and doing the right thing *for the right reasons*. Doing the right thing for the right reasons makes the action a morally worthy one, as it exhibits a good quality of the will. It is possible for a theory to hold that the ‘good will’ is one that understands the underlying justification of an action, but that seems overly demanding. If consequentialism is the correct theory, then demanding that people must explicitly act intentionally to maximize the good would result in fewer morally worthy actions than seems plausible. The ‘for the right reasons’ must be understood as allowing for no explicit invocation of the true justifying standard.

This has led to the development of theories that advocate indirection. First, we need to distinguish two ways that indirection figures into moral philosophy.

1. Indirection in evaluation of right action.
2. Indirection in that the theory does not necessarily advocate the necessity of aiming for the right action.

To use Utilitarianism as an example again, Rule Utilitarianism is an example of the first sort of indirection (Hooker 2000), *Sophisticated Consequentialism* is an example of the second sort of indirection (Railton 1984). One might hold that some versions of Aristotelian Virtue ethics, such as Rosalind Hursthouse’s version, also are of the first type, since right action is understood in terms of virtue. One could imagine an indirect consequentialist view with a similar structure: the right action is the action that the virtuous person would perform, where virtue is understood as a trait conducive to the good, instead of by appeal to an Aristotelian notion of human flourishing.

The second sort relies on the standard/decision-procedure distinction. Railton argues that personal relationships are good for people, and

explicitly trying to maximize the good is not a part of our relationship norms, so it is likely good that we develop dispositions to focus on and pay special attention to our loved ones. The account is open to the possibility that people who don’t believe in consequentialism have another way of deciding how to act that is correlated with promotion of the good. If the criteria a theory sets out need not be fulfilled by the agent guiding herself with the reasons set out by the criteria, then it is termed *self-effacing*. When a theory is self-effacing, it has the problem of alienating a person from the justification of her own actions. A middle ground, which is closer to Railton’s view, holds that the correct justification is a kind of “touchstone” to the morally good person – consulted periodically for self-regulation, but not taken explicitly into consideration in our ordinary, day-to-day lives. In this way, the theory would not be utterly self-effacing and the agent would still understand the moral basis for her own actions.

## Bibliography

- Alicke, Mark, David Rose and Dori Bloom, 2011, “Causation, Norm Violation, and Culpable Control,” *Journal of Philosophy*, 108(12): 670–696.
- Annas, Julia, 2011, *Intelligent Virtue*, New York: Oxford University Press.
- Arpaly, Nomy, 2002, *Unprincipled Virtue*, New York: Oxford University Press.
- Baier, Annette, 1985, *Postures of the Mind*, Minneapolis, MN: University of Minnesota Press.
- Baron, Marcia, 1991, “Impartiality and Friendship,” *Ethics*, 101(4): 836–857.
- , 1995, *Kantian Ethics Almost Without Apology*, Ithaca, NY: Cornell University Press.
- Clarke, Stanley G, 1987, “Anti-Theory in Ethics,” *American Philosophical Quarterly*, 24(3): 237–244.

- D'Arms, Justin and Daniel Jacobson, 2000, "The Moralistic Fallacy: On the 'Appropriateness' of Emotions," *Philosophy and Phenomenological Research*, 61(1): 65–90.
- Darwall, Stephen, 2006, *The Second-Person Standpoint*, Cambridge, MA: Harvard University Press.
- Dreier, Jamie, 1993, "Structures of Normative Theories," *The Monist*, 76(1): 22–40.
- Driver, Julia, 2012, "What the Objective Standard is Good For," in Mark Timmons (ed.), *Oxford Studies in Normative Ethics*, New York: Oxford University Press, 28–44.
- Dworkin, Ronald, 1977, *Taking Rights Seriously*, Cambridge, MA: Harvard University Press.
- Foot, Philippa, 1967, "Abortion and the Doctrine of Double Effect," *Oxford Review*, 5: 5–15.
- Graham, Peter, 2010, "In Defense of Objectivism About Moral Obligation," *Ethics*, 121(1): 88–115.
- Hare, R. M., 1965, *Freedom and Reason*, New York: Oxford University Press.
- Herman, Barbara, 1985, "The Practice of Moral Judgment," *Journal of Philosophy*, 82(8): 414–436.
- Hill, jr., Thomas E., 1987, "The Importance of Autonomy," in Eva Kittay and Diana Meyers (ed.) *Women and Moral Theory*, Totowa, NJ: Rowman & Allanheld, 129–138.
- Hooker, Brad, 2000, *Ideal Code, Real World*, New York: Oxford University Press.
- Hurka, Thomas, 2001, *Virtue, Vice, and Value*, New York: Oxford University Press.
- Hursthouse, Rosalind, 1999, *On Virtue Ethics*, New York: Oxford University Press.
- Jackson, Frank, 1991, "Decision-theoretic Consequentialism and the Nearest and Dearest Objection," *Ethics*, 101(3): 461–482.





- Jeske, Diane, 2008, *Rationality and Moral Theory: How Intimacy Generates Reasons*, New York: Routledge.
- Joyce, Richard, 2001, *The Myth of Morality*, New York: Cambridge University Press.
- Keas, Michael, 2018, "Systematizing the Theoretical Virtues," *Synthese*, 195: 2761–2793.
- Kagan, Shelley, 1989, *The Limits of Morality*, New York: Oxford University Press.
- Kamm, Frances, 2007, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*, New York: Oxford University Press.
- Kant, Immanuel, 1785 [2012], *Groundwork of the Metaphysics of Morals*, tr. by Mary Gregor and Jens Timmerman, New York: Cambridge University Press, 2012.
- Knobe, Joshua, 2003, "Intentional Action in Folk Psychology: An Experimental Investigation," *Philosophical Psychology*, 16(2): 309–325.
- Kripke, Saul, 1982, *Wittgenstein on Rules and Private Languages*, Cambridge, MA: Harvard University Press.
- Louden, Robert, 1990, "Virtue Ethics and Anti-Theory," *Philosophia*, 20(1–2): 93–114.
- Markovits, Julia, 2014, *Moral Reason*, New York: Oxford University Press.
- Mason, Elinor, 2003, "Consequentialism and the 'Ought Implies Can' Principle," *American Philosophical Quarterly*, 40(4): 319–331.
- , 2019, *Ways to Be Blameworthy: Rightness, Wrongness, and Responsibility*, New York: Oxford University Press.
- McDowell, John, 1979, "Virtue and Reason," *The Monist*, 62(3): 331–350.
- Moody-Adams, Michelle, 2002, *Fieldwork in Familiar Places: Morality, Culture, and Philosophy*, Cambridge, MA: Harvard University Press.
- Moore, G. E., 1903 [1993], *Principia Ethica*, ed. Thomas Baldwin, New York: Cambridge University Press, 1993.

- Nagel, Thomas, 1979, "The Fragmentation of Value," in *Mortal Questions*, New York: Cambridge University Press, 128–141.
- Nietzsche, Friedrich, 1887 [1998], *On the Genealogy of Morality*, Maudemarie Clark and Alan J. Swensen (trans.), Indianapolis, IN: Hackett Publishing.
- Norcross, Alastair, 2020, *Morality By Degrees*, New York: Oxford University Press.
- Olson, Jonas, 2004, "Buck-Passing and the Wrong Kind of Reasons," *Philosophical Quarterly*, 54(215): 295–300.
- Parfit, Derek, 1984, *Reasons and Persons*, Oxford: Clarendon Press.
- Pettit, Phillip, 1997, "The Consequentialist Perspective," in *The Three Methods of Ethics*, by Marcia Baron, Phillip Pettit, and Michael Slote, Oxford: Blackwell, 92–174.
- Pettit, Phillip, and Michael Smith, 2000, "Global Consequentialism," in Brad Hooker, et al. (eds.), *Morality, Rules, and Consequences*, Edinburgh: University of Edinburgh Press, 121–133.
- Phillips, David, 2019, *Rossian Ethics*, New York: Oxford University Press.
- Piper, Adrian, 1987, "Moral Theory and Moral Alienation," *Journal of Philosophy*, 82(2): 102–118.
- Portmore, Douglas, 2011, *Commonsense Consequentialism*, New York: Oxford University Press.
- Rabinowicz, Wlodek and Toni Ronnow-Rasmussen, 2004, "The Strike of the Demon: on Fitting Pro-Attitudes and Value," *Ethics*, 114(3): 391–423.
- Railton, Peter, 1984, "Alienation, Consequentialism, and the Demands of Morality," *Philosophy and Public Affairs*, 13(2): 134–171.
- Rawls, John, 1971, *A Theory of Justice*, Cambridge, MA: Belknap Press.
- Scanlon, T. M., 1998, *What We Owe to Each Other*, Cambridge, MA: Belknap Press.
- , 2008, *Moral Dimensions*, Cambridge, MA: Harvard University Press.

- Scheffler, Samuel, 1982, *The Rejection of Consequentialism*, New York: Oxford University Press.
- Schneewind, J. B., 1963, "First Principles and Common-sense Morality in Sidgwick's Ethics," *Archiv für Geschichte der Philosophie*, 45(2): 137–156.
- , 1990, "The Misfortunes of Virtue," *Ethics*, 101(1): 42–63.
- Schofield, Paul, 2021, *Duty to Self: Moral, Political, and Legal Self-Relation*, New York: Oxford University Press.
- Sen, Amartya, 2000, "Consequential Evaluation and Practical Reason," *The Journal of Philosophy*, 47(9): 477–502.
- Sidgwick, Henry, 1874 [1907], *The Methods of Ethics*, London: Macmillan. [The seventh edition was published in 1907.]
- Singer, Marcus, 1986, "Ethics and Common Sense," *Revue Internationale de Philosophie*, 40(158): 221–258.
- Slote, Michael, 1985, *Common-Sense Morality and Consequentialism*, New York: Routledge & Kegan Paul.
- , 2007, *The Ethics of Care and Empathy*, New York: Routledge.
- Smith, Holly, 2018, *Making Morality Work*, New York: Oxford University Press.
- Smith, Michael, 2009, "Two Kinds of Consequentialism," *Philosophical Issues*, 19(1): 257–272.
- Stark, Cynthia, 1997, "Decision Procedures, Standards of Rightness and Impartiality," *Noûs*, 31(4): 478–495.
- Stocker, Michael, 1976, "The Schizophrenia of Modern Ethical Theories," *Journal of Philosophy*, 73(14): 453–466.
- Strawson, Peter, 1961, "Social Morality and Individual Ideal," *Philosophy*, 36(136): 1–17.
- Street, Sharon, 2006, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies*, 127(1): 109–166.
- Thomson, Judith Jarvis, 1976, "Killing, Letting Die, and the Trolley Problem," *The Monist*, 59(2): 204–217.

- Wiland, Eric J, “The Incoherence Objection in Moral Theory,” *Acta Analytica*, 25(3): 279–284.
- Williams, Bernard, 1985, *Ethics and the Limits of Philosophy*, New York: Oxford University Press.
- Wolf, Susan, 1982, “Moral Saints,” *Journal of Philosophy*, 79(8): 419–439.
- , 2014, “Loving Attention: Lessons in Love from *The Philadelphia Story*,” in Susan Wolf and Christopher Grau (eds.), *Understanding Love: Philosophy, Film, and Fiction*, Oxford: Oxford University Press, pp. 369–386.
- Zagzebski, Linda Trinkhaus, 2017, *Exemplarist Moral Theory*, New York: Oxford University Press.

## Academic Tools

-  How to cite this entry.
-  Preview the PDF version of this entry at the Friends of the SEP Society.
-  Look up topics and thinkers related to this entry 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: deontological | ethics: virtue | morality, definition of | value theory

## Notes to Moral Theory

1. There is a more refined use of “Common-sense Ethics” associated with Thomas Reid who was responding to what he viewed as the moral skepticism espoused by David Hume.
2. See Sidgwick 1874 [1907] and Schneewind 1963.

Copyright © 2022 by the author  
Julia Driver