https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

MORAL CONSTRUCTION AS A TASK: SOURCES AND LIMITS

By Thomas E. Hill, Jr.

I. Preliminary Remarks, Aims, and Preview

Many different questions have been debated under the rubrics of "objectivity" and "relativity" of morals. My plan is to begin by distinguishing several of these questions as background discussion. Then, after a few remarks about the earlier questions, I sketch a broadly Kantian position on the last two of these questions: *First,* how, if at all, can we derive, justify, or support specific moral principles and judgments from more fundamental moral standards and values? *Second,* how, if at all, can (alleged) fundamental standards be defended? Famously, Kant attempted to provide nonskeptical answers to both questions. His explicit answers and arguments have often been criticized, and there are probably few today who accept them all without qualification. Many contemporary philosophers, however, have been influenced by his ideas about the nature and structure of moral thinking, his attempt to articulate a supreme moral standard, and his strategy for defending moral principles. My sketch here of a broadly Kantian perspective represents my attempt to develop and extend some of these ideas beyond Kant's texts.[1] The influence of John Rawls's theory of justice may also be evident.

To preview, the main features of my broadly Kantian position on our two main questions are these. *First,* Kant's later formulations of the Categorical Imperative provide important elements of a deliberative procedure that can be used to guide and constrain our thinking about more specific moral principles in a way that helps to correct for mere personal

[1] By calling a perspective "broadly Kantian" I mean to acknowledge that it draws heavily from Kant's ethical writings but does not follow Kant's texts in all respects. The broadly Kantian perspective that I sketch here and elsewhere modifies and supplements Kant's theory in various ways, one of which is its attempt to disassociate normative aspects of the theory from Kant's theory of noumenal freedom (as this is usually understood). I give a fuller (though still incomplete) description of the broadly Kantian perspective I have in mind in other essays, especially Thomas E. Hill, Jr., *Respect, Pluralism, and Justice: Kantian Perspectives* (Oxford: Oxford University Press, 2000), chaps. 2, 4, and 8; Hill, *Human Welfare and Moral Worth: Kantian Perspectives* (Oxford: Oxford University Press, 2002), chap. 3; and Hill, "Assessing Moral Rules: Utilitarian and Kantian Perspectives," *Normativity, Philosophical Issues (A Supplement to Noûs),* vol. 15, edited by Ernest Sosa and Enrique Villanueva (2005): 158–78.

DOI: 10.1017/S0265052508080084

© 2008 Social Philosophy & Policy Foundation. Printed in the USA.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

preference and parochial bias.[2] The procedure is far less than a determinate decision procedure or strict "derivation" of particular moral conclusions, but it promises a way of supporting judgments that initially may have appeared more doubtful. *Second,* Kant has two strategies for defending the ideas in his various formulations of the Categorical Imperative. The first has merit, at least when modified and presented more modestly as akin in some respects to Rawls's method of reflective equilibrium. The background of this strategy, too frequently ignored, is Kant's principled rejection of previous errors in foundational moral theory. Also, though not a major focus here, Kant's second line of defense, at least in its final modified form, may still serve its purpose in practical, rather than metaphysical, reflection and dialogue. If what is questioned is the rational authority of the basic standards identified by the previous method, the best we can do, and often all that is needed, may be to appeal to deep pervasive commitments and consistency in the practice and serious thinking of each rational moral agent.

Words such as "construction," "objectivity," and "relativity" are at best blunt instruments for discussion, and by now they have been used in so many ways they may have outlived their usefulness. Nevertheless, to connect my discussion with some recent literature, I propose that the broadly Kantian perspective sketched here leaves us with two ways in which moral *construction* is an ongoing *task.* First, and most obviously, *defining the broadly Kantian deliberative perspective* is an incomplete work in progress—an attempt to articulate for practical purposes an appropriate framework for deliberating about the interpretation and limits of specific moral principles (or "rules"). This construction project starts from the assumption that Kant's own efforts to articulate a comprehensive supreme moral principle were inadequate for the purpose but contain important, promising elements. The aim would be to put these Kantian elements together, modifying and supplementing them as needed in the light of arguments, to build an account of an appropriate deliberative perspective or stance for addressing more specific moral issues. The challenge would be to identify and defend certain more basic values as constituting the framework for deliberation about further, often more controversial issues, for example, the interpretation and limits of rules about lying, preemptive killing, and torture. To continue the metaphor of construction, this *substantive deliberation about more specific principles* can be seen as a second, more practical construction project. In any broadly Kantian view, controversies about the interpretation of (and exceptions to) specific moral rules cannot be resolved by science, metaphysics, theology, or appeal to any authority based merely on power, tradition, or convention. Instead, such

[2] Immanuel Kant, *Groundwork for the Metaphysics of Morals,* trans. Arnulf Zweig, ed. Thomas E. Hill, Jr., and Arnulf Zweig (Oxford: Oxford University Press, 2002), 222–37 [4:421–38]. Bracketed numbers in all citations of Kant's works refer to volume and page numbers in the standard Prussian Academy edition.

principles and their limits must, in a sense, be "constructed" in reflection from an appropriate moral point of view by the moral agents who must be subject to them.[3]

My remaining discussion will be divided into the following three main sections: In Section II, as background, I distinguish several different questions that arise in debates about moral relativism and objectivity. Then, in Section III, I sketch my broadly Kantian perspective for assessing specific moral principles. Finally, in Section IV, I describe briefly Kant's grounds for rejecting previous moral theories and his two strategies for defending a proposed deliberative perspective.

II. Different Questions Regarding Relativism and Objectivity

Some worry and others happily accept that moral principles and judgments may be "relative" and "subjective," as opposed to "absolute" and "objective." These terms, unfortunately, have been understood in so many different ways that they are now more often confusing than helpful toward understanding deep philosophical concerns.[4] At least their use in each context needs to be explained and supplemented with discussion in other terms. So I shall try to distinguish several questions without *relying* on the

[3] The idea that moral principles can be *constructed* is a metaphor that can be illuminating in some respects but is also liable to be misleading. For example, the metaphor is intended to suggest that in order to reach reasonable conclusions about what moral principles require and what exceptions they allow, what is needed for practical purposes is not acute "perception" of natural or nonnatural metaphysical facts but rather reflection and discussion about what standards we have most reason to endorse in the light of certain basic moral values, norms of rationality, and relevant facts about the human condition and the more local conditions in which the principles are to be applied. The metaphor of construction would be misleading if it suggested that we may construct whatever principles we please as if we were children given a pile of miscellaneous materials with which to build something but without any given blueprint, purpose, or standard for construction. When I suggest that we may think of *substantive principles* (for example, about lying, stealing, and killing) as *constructed* (or to be constructed) from a broadly Kantian deliberative perspective, I do not mean to imply that they are (or are to be) constructed from morally neutral or value-free assumptions or from values that are themselves constructed in the same sense. In saying, in addition, that we may think of *articulating the broadly Kantian deliberative perspective* as a *task for construction,* I mean to suggest only that it remains an unfinished task in moral theory to work out and express as clearly as possible a description of the most basic constraints and guiding values that are reasonable, and appropriately called (broadly) "Kantian," as a framework for assessing more substantive and derivative moral principles. When understood this way, construction of substantive principles need not be redundant or viciously circular, and construction of the deliberative basic perspective does not necessarily presuppose any particular metaethical theory. A more radical and controversial claim would be that *all* moral values and principles, including those treated as basic in a broadly Kantian deliberative perspective, are *constructed* in some further sense.

[4] Philosophers have often noted important distinctions among the questions relevant to discussions of relativism and objectivity. An excellent example is Richard Brandt, *Ethical Theory* (Englewood Cliffs, NJ: Prentice Hall, 1959), chap. 11.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

terms "relative," "absolute," "subjective," and "objective," even though I will use some of these terms as temporary *labels* for ideas that I explain.

A. Linguistic relativity and related deeper disputes

Some words are incomplete without explicit or implicit reference to something further. An obvious example is "relevant." Back in the 1960s, students often clamored for new courses that were "relevant" but without specifying what they were supposed to be relevant *to.* Some no doubt had the general idea that new courses should be relevant to their interests in stopping the war in Vietnam, reforming or destroying the existing government, living untraditional lifestyles, etc., but others seemed to think that the word "relevant" by itself expressed a complete idea. More commonly, people often talk about various things being *important* without saying *to whom* or *for what* they are significant. No doubt many could pause, if asked, and fill in the assumed or neglected qualification, but too often conversations proceed as if no further comment is required, some things just being important period, not *for* any specifiable purpose and not *to* you, or to me, or even to all human beings, to all rational beings, to all living things, etc. In these cases, we naturally wonder what the standards for evaluating such claims are and whether, after all, the claims make sense.

Thomas Hobbes says that "good" is always "relative to the person," usually relative to the person that uses the word but relative to the sovereign authority in a state when the sovereign makes a ruling about what is to be regarded good.[5] At least part of the point seems to be that the word "good," like "relative" and "important," is incomplete by itself, and to make sense needs to be supplemented, for example, by specifying to or for whom, or from whose perspective, something is alleged to be good. We might add a variety of other possible supplements, such as "good for . . . (a purpose)," "a good . . . (kind of thing)," or "good from the point of view of . . . (a group, project, or institution with shared aims, etc.)." Similar things might be said of "valuable," "desirable," "ideal," and other value terms (including negative terms such as "bad," "undesirable," "repugnant," etc.). If we want a label, we might use *linguistic relativity* for this idea that familiar value terms, such as "good," are incomplete without further specification.

Underlying any dispute about the linguistic point, however, is likely to be a deeper disagreement in moral theory or metaphysics. Plato held that there is a Form of Goodness that exists independently of existing earthly beings, their purposes, and their natural kinds. Hobbes made his linguis-

[5] Thomas Hobbes, *Leviathan*, ed. Edwin Curley (Indianapolis, IN: Hackett Publishing Co., 1994), part I, chap. 6, pp. 28–29. This famous passage in Hobbes is open to several different interpretations. My concern here is with the idea I draw from Hobbes, which may or may not be exactly what Hobbes meant.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

tic point in a way that also expressed his antipathy to a Platonic conception of goodness. Unlike Plato, he thought that we call things *good* because we desire them, not that we desire them because (independently) they are good. Like Plato, G. E. Moore also implied that various things are good in themselves, or good *period,* as opposed to good *for* something, good *to* someone, good *of a kind,* etc. His underlying point was not merely linguistic but a controversial metaphysical thesis about the existence of goodness as a real nonrelational, nonnatural property.[6] Disputes about this go beyond linguistics and more properly fall under later questions about basic standards and justification.

B. Causal dependence on culture and history

When discussing whether morality is "relative," one often hears the claim that moral beliefs are determined by cultural and historical factors that vary widely. This is an empirical thesis about the causes of moral beliefs. Its relevance to disputes about the truth or validity of universal moral principles is less clear than is often supposed. There are two main points to consider. First, if true, the thesis would suggest (but does not entail) that there are no moral principles that are universally *believed.* Although it is not necessary, it is perhaps natural to expect that variable causes will produce variable effects. So the causal dependence claim is significant partly because it can be taken to support a "relativist" position on the next issue on my list, that is, whether or not all moral *beliefs* vary with cultural and historical conditions. The second and more important point to consider, however, is that the causal dependence of moral beliefs on variable cultural and historical factors is thought to undermine moral theories committed to other sources of moral belief—intuition, conscience, innate reason, divine commands, or universally shared sentiments. It is a fallacy to argue that a belief must be false because it was caused by something other than evidence for its truth. Hence, tracing the causes of moral beliefs to variable cultural conditions unrelated to moral theorists' alleged sources does not prove that the normative principles of these theorists are false, but it tends to undermine moral theories that depend on alternative causal stories to explain how they know their principles are true or valid.

C. Diversity of many or all moral beliefs

As I just noted, sometimes the "relativist" position is an empirical thesis that what is *accepted* as morally right or good varies widely over times and places. A radical version would hold that there are *no* universally shared moral beliefs or values, and for every moral belief and value

[6] G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903, 1959).

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

accepted in one culture there are *incompatible* beliefs and values in some other culture. That many particular moral beliefs diverge (for example, beliefs regarding sex, property, or the conduct of war) is undeniable, but it remains controversial whether there are some general moral values underlying the diversity of particular moral beliefs accepted in different cultures. This is an empirical question that cannot be easily settled. (A complication is that to assess evidence one must first take a position about what beliefs and norms count as *moral*.) When there is cultural variance in beliefs about facts such as the shape of the earth, there are ways to resolve the issue, showing that certain beliefs are true (the earth is round) and others mistaken (the earth is flat). Mere diversity of belief does not entail that there is no fact of the matter. The claim that all moral beliefs vary with cultural and historical conditions, then, does not entail that there are no true or valid universal moral beliefs. It does, however, raise doubts about moral theories committed to certain explanations of moral beliefs that would lead us to expect almost universal agreement about moral issues. If, for example, the theorists tell us that learning moral truths is like perception of colors or "seeing" simple mathematical propositions to be "self-evident," then they will be hard-pressed to explain pervasive and persistent diversity of moral beliefs.

D. Normative dependence of specific moral rules and particular judgments on context

Some believe that there are fairly specific, substantive moral rules that are true or valid in all human conditions. Many understand the Ten Commandments this way. In a notorious late essay, Immanuel Kant argues that telling lies is always wrong, and even some who espouse "virtue ethics" hold that acts of certain kinds (such as torture or intentionally killing innocent persons) are always wrong.[7] By contrast, act-utilitarians maintain that specific conclusions about right and wrong depend on whatever will have (or is expected to have) the best consequences in the context. They typically affirm this act-utilitarian standard as true or valid but grant that the rightness or wrongness of specific acts can vary with the circumstances. For example, according to the act-utilitarian standard, malicious self-serving lies tend to be wrong because they cause more harm than good, though it would not have been wrong to tell a lie to a Nazi at the door to save the Jewish family hiding in one's attic. Even

[7] For Kant's extreme position on lying, see "On a Supposed Right to Lie from Philanthropy," in Immanuel Kant, *Practical Philosophy,* trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1996), 605–15 [8:423–30]. Elizabeth Anscombe's view is expressed in her classic essay "Modern Moral Philosophy," *Philosophy* 33 (1958): 1–19. Some other philosophers associated with virtue ethics (for example, Philippa Foot) have reportedly endorsed this view in discussions.

deontologists such as W. D. Ross hold that, although there are true universal principles of *prima facie* duty, particular judgments about one's "actual duty" in a situation depend also on the variable facts of the case at hand.[8] Contemporary moral "particularists" (for example, Jonathan Dancy) go even further, maintaining that correct moral judgments are determined by the particular facts of each case independently of moral principles.[9] Dancy does not deny that *particular* moral judgments are true, "objective," and supported by reasons, but only that there are valid and useful moral *generalizations.* All of these theorists hold that there are true or valid particular moral judgments, as determined either by universal moral principles or by good reasons in a context. Their primary disagreement is about the *scope* of intermediate-level moral rules or principles—for example, those forbidding lying, promise-breaking, torturing prisoners, and engaging in extramarital sex.[10] Are these valid rules or true generalizations for everyone, or only within certain cultures, or only for certain groups or individuals, or not at all?

E. Questions about the scope and authority of basic standards and deliberative procedures

Those who hold that specific moral rules and judgments are true or false, or valid or invalid, depending on their context typically believe that there are more basic, general standards or procedures that explain this variability of specific rules and judgments. Both act-utilitarians and rule-utilitarians affirm such basic standards, and so do deontologists such as

[8] W. D. Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930), chap. 2. Ross held that it is the principles of prima facie duty together with the facts of the particular case at hand that determine one's actual duty, but because none of those principles is absolute and they cannot be ranked with respect to one another, one's actual duty can vary with even slight changes in circumstance. To say that one has *a prima facie duty* to keep one's promise in a particular case is, in effect, to say that the fact that one promised is a *pro tanto* or defeasible moral reason to keep the promise or, in other words, a reason for doing what was promised that is sufficient in the absence of weightier countervailing moral reasons but not necessarily decisive in all contexts.

[9] Jonathan Dancy, *Ethics without Principles* (Oxford: Oxford University Press, 2004), chap. 1.

[10] Principles can be more or less general and comprehensive. In Kant's moral theory, the various formulations of the Categorical Imperative are supposed to express the supreme moral principle, which is the most general principle relevant to all questions about what we morally ought to do. What I call "intermediate-level principles" (or sometimes just "specific principles") are not as general and comprehensive, applying only to certain kinds of often-recurring cases, for example, cases involving lies, theft, various sexual practices, developing one's talents, giving aid to the needy, etc. These are intermediate in generality between the Categorical Imperative and extremely specific rules and judgments concerned with a narrowly restricted range of cases (e.g., the special responsibilities of doctors and social workers regarding patients with Alzheimer's disease). Kant outlines a system of intermediate-level principles for law and individual ethics in Immanuel Kant, *The Metaphysics of Morals,* trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1996).

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

Ross and Kant.[11] The standards or procedures are commonly regarded as true or valid for all competent adults. They are conceived as the ultimate test of what is moral, and for many they are also (or thereby) an ultimate rational authority. Opposing theorists (and some who are opposed to all theory) deny that there are universally valid basic moral standards or deliberative procedures that can guide and constrain our assessment of more particular moral rules and judgments (or make some true and others false).[12] There are different ideas, of course, about what it means for basic standards to be true, valid, and rational.

F. The possibility of proof, justification, or defense of basic standards

Many who accept that there are true or valid basic moral standards offer arguments of various kinds in support of their position, but some simply pronounce their standards to be self-evident. Kant and John Stuart Mill are two prominent moral philosophers who took seriously the tasks of explaining a sort of "proof" or defense that they believed possible and of offering arguments to serve the purpose. There are different views, of course, about what would constitute a proof, justification, or defense, and many theorists are skeptical about whether these are possible in any form. What constitutes a proof, justification, or defense depends heavily on how a given moral theory construes moral judgments—as stating natural facts, expressing the speaker's attitudes, ascribing real nonnatural properties to things, applying necessary rational principles, and so on.

There are, no doubt, other questions and further distinctions in disputes about moral relativity and objectivity, but my aim here has been only to set the stage for discussion. The primary focus of the rest of this essay will be on the last two issues—more specifically, on a broadly Kantian position on *whether there are general and basic moral standards* that determine which more specific moral principles and judgments are justifiable and *how, if at all, the more basic standards can be defended.* These are

[11] The act-utilitarian standard (roughly) is that one ought to do whatever brings about, or is expected to bring about, the most good (e.g., pleasure or well-being) in the long run. The standards of rule-utilitarianism are more complex to state because of various necessary qualifications, but the main idea is that one ought to conform to a set of rules (a code) the internalization of which by most people would bring about the most good. See Brad Hooker, *Ideal Code and Real World* (Oxford: Clarendon Press, 2000), 1–2 and 33. Ross's basic standards are his principles of prima facie duty concerning fidelity, gratitude, reparation, justice, self-improvement, non-injury, and beneficence. See note 8. Kant's basic standards or deliberative procedures are expressed in the several formulations of the Categorical Imperative. See note 2.

[12] The opposing theorists include not only nihilists (who dismiss morality entirely) but also subtle expressivists (who deny literal *truth* and validity to moral principles but do not deny their importance) and particularists (who affirm particular moral truths, possibly even similar ones in different cultures, but deny that there are basic *general* standards). Another possible view would be that there are general standards that are basic and true or valid within different historical or cultural contexts but no basic moral standards that are true or valid for all contexts.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

difficult questions because they require not only interpreting and extending Kant's ideas but also discussing them in the terms of our contemporary debates in moral theory.

Before turning to the main two questions, however, here are a few brief comments on my broadly Kantian perspective on the other questions—comments which may explain why those are not my focus of concern. The questions in subsections B and C above are entirely matters for empirical investigation. Kant had opinions about them but no relevant expertise, and the main features of his moral theory, arguably, do not depend on his opinions on such empirical questions.[13] The main question in subsection A presumably belongs to traditional metaethics because it concerns the meaning, or proper use, of the word "good" and related terms. Kant's position on this matter could easily be misunderstood. He repeatedly writes

[13] The relevance of various empirical facts to Kant's basic normative claims is not obvious. Kant evidently believed that virtually every mature and competent moral agent, at least every one in relatively "advanced" or "enlightened" parts of the world (and history), has practical reason and a conscience that acknowledges "the moral law," no matter how badly he or she may behave. He also thought that reason can be latent and conscience dulled, and that "evil" people intentionally elect to subordinate their moral predispositions. See Kant, *The Metaphysics of Morals,* 160–61 [6:400–401] and 189–91 [6:438–40], and Kant, *Religion within the Boundaries of Mere Reason,* ed. Allen Wood and George di Giovanni (Cambridge: Cambridge University Press), 179–80 [6:184–87] and 45–65 [6:19–44]. Kant's deplorable anthropological remarks about peoples of other "races" at least suggest that, in his opinion, even the capacity for morality may be more limited than his inspiring words about human dignity would seem to imply. See Bernard Boxill and Thomas E. Hill, Jr., "Kant and Race," in *Race and Racism,* ed. Bernard Boxill (Oxford: Oxford University Press, 2001), 448–71. It remains a difficult question to what extent Kant's main theses and arguments depend on his empirical beliefs and prejudices. *Groundwork* I starts with assertions about common rational knowledge of morality and proceeds to analyze their presuppositions. If Kant was mistaken about how wide the community that shared the "common" beliefs is, this would limit the interest in his main argument, making it applicable only to those who share the common beliefs. This, however, would not necessarily refute or render useless the sort of conditional ("analytical" style) argument that he admitted the argument in *Groundwork* I to be. Similarly, *Groundwork* II argues that our common conception of duty has certain striking presuppositions, but Kant concedes rather dramatically that this does not entail even that "we" who have that conception are really under moral obligations—and, all the more, the argument clearly does not entail that human beings who lack the conception are under moral obligations. In *Groundwork* III, and the *Critique of Practical Reason,* Kant tries to convince readers of the stronger claim that all agents with practical reason are subject to the moral requirements prescribed by the Categorical Imperative, but he is committed to the idea that these arguments cannot rest on empirical claims (for example, the claim that, as a matter of fact, all cognitively competent human beings acknowledge these moral requirements or would do so after being given rational arguments). See sections I and II of Kant's *Groundwork* and, for an interpretation of the main arguments of those sections, see my "Analysis of Arguments," in Immanuel Kant, *Groundwork for the Metaphysics of Morals,* trans. Arnulf Zweig and ed. Thomas E. Hill, Jr., and Arnulf Zweig (Oxford: Oxford University Press, 2002), 195–245 [4:393–445] and 109–14. See also Immanuel Kant, *Critique of Practical Reason,* trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1997), 28–64 [5:30–50]. If Kant's conception of duty and a good will were limited to eighteenth-century Prussians, the interest in his arguments would be severely limited in scope, but there is apparently a wider audience that shares Kant's conception of duty, or at least the belief that a moral duty is what one has sufficient reason to do and not just because it serves one's interests. Whether, despite objections, Kant's arguments have any merit remains a controversial question, not easily settled by reference to anthropological facts and still subject to doubts of other kinds.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

of moral acts as "good in themselves" and thus does not buy into Hobbes's thesis that "good" is always "relative to the person" (more specifically, to the desires of the speaker or the command of the sovereign). In Kant's view, however, there is nothing, not even what he calls "good in itself," that is "good" in a practical sense independent of its relation to actual or possible rational agents and their reasons for intending, willing, and acting. Kant does not believe in intrinsic values in the metaphysical sense that, for example, is central in G. E. Moore's intuitionist ethics.

Regarding the issue discussed in subsection D, Kant famously held that there are some specific moral principles that hold for all people at all times. His notorious example was an absolute prohibition on (serious) lying, but it is noteworthy that even famous advocates of "virtue ethics" (for example, Elizabeth Anscombe) apparently accept some completely inflexible universal principles about other matters, such as the intentional killing of innocent people.[14] On my broadly Kantian view, when (if ever) principles regarding lying and killing innocents are subject to exceptions must be determined by reflection from a more basic deliberative perspective. What this is and whether it is defensible—the questions in subsections E and F—are the topics for my remaining discussion.

III. A Broadly Kantian Perspective for Assessing Specific Moral Principles

Here I can only give a brief summary of relevant points in the Kantian perspective that I have in mind. Kantian moral theory belongs to a rationalistic tradition that maintains that the most fundamental moral values and standards are accessible through the use of practical reason, applicable to all rational human beings, and contrary to reason to dismiss or violate.[15] These fundamental values and standards are supposed to be expressed in Kant's several formulations of the Categorical Imperative and accompanying explanations.[16] Kant regarded these basic standards as presupposed by, and even implicit in, ordinary moral deliberation and judgment. He illustrates the guiding function of two versions of the Categorical Imperative by applying them directly to four examples that he

[14] See note 7 above.

[15] In Kantian theory, as I understand it, "values" are not entities or facts independent of what rational agents would care about (or "value") in various conditions. To say that humanity is a basic Kantian value, then, is not to claim that it has a perceptible quality that explains and justifies the idea that we have a rational requirement or good reasons to respect it. Saying that humanity is a basic value implies that it should be respected; i.e., we have a rational requirement or good reasons to respect it, because these are just different ways of saying the same thing. The Kantian view, so construed, is an instance of what T. M. Scanlon calls the "buck-passing" view of value—that is, value-claims do not state the *justifying basis* of reason-claims but rather *express* the idea *that there are* good reasons for taking certain (not yet specified) attitudes and actions toward the thing. T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998), 95–100.

[16] Kant, *Groundwork for the Metaphysics of Morals,* 222–37 [4:421–38].

sketches. In his later, more thorough treatment of practical problems, however, he appeals to basic standards together with general facts about the human condition to identify and (sometimes) argue for a system of general principles about property, marriage, legal authority, punishment, lying, self-neglect, helping others, gratitude, respect, and friendship.[17] This is a system of principles of mid-level generality, standing between the most comprehensive supreme moral principle (expressed in the Categorical Imperative) and moral judgments in particular contexts. Some principles state *juridical duties,* which can and ought to be enforced by state authorities. Other principles state directly *ethical duties,* which require us to adopt maxims of various sorts to promote the ends of self-perfection and the happiness of others. The strict *perfect duties* have the form "Never . . ." or "Always . . . ," but the *imperfect duty* to help others leaves room for judgment and choice regarding when, how, and to what extent to help.

According to Kantian theory, then, there is a structure that can guide us to correct moral judgments in particular cases. The ultimate standard, a principle or procedure expressed as the Categorical Imperative, is taken to be a comprehensive, fundamental, and even necessary requirement of practical reason—and so the standard for our reasoning about what we ought to do. Applying this standard to pervasive human conditions, Kant attempts to identify (and justify relative to the standard) a system of mid-level principles of several kinds intended to be applicable independently of variations in local conditions. The principles of perfect duty must never entail contradictory prescriptions, but the "grounds of duty" from which they are drawn may be in tension.[18] All the more, imperfect duties may be in tension, for promoting one moral end may limit the ways that one can pursue another. The system is supposed to be structured and consistent, but its imperfect duties allow considerable room for choice, and all duties require interpretation and judgment appropriate to the context.[19]

When explaining the basic Kantian standard for identifying more specific moral principles, most contemporary reconstructions and developments of Kant's moral theory concentrate on interpretations of either Kant's formula of universal law or his formula of humanity as an end in itself.[20] Skeptical of these (or any formulas) as case-by-case moral deci-

[17] Kant, *The Metaphysics of Morals,* parts I and II, 1–170 [6:203–413].
[18] Ibid., 16–17 [6:224–25].
[19] Ibid., 153 [6:390].
[20] The formula of universal law states: "Act only on that maxim by which you can at the same time will that it should become a universal law." Kant, *Groundwork for the Metaphysics of Morals,* 222 [4:421]. A variation that Kant uses in discussing his examples is: "Act as though the maxim of your action were to become by your will a universal law of nature." Kant, *Groundwork,* 222 [4:421]. The formula of humanity states: "Act in such a way that you treat humanity, whether in your own person or in any other person, always at the same time as an end, never as a means." Kant, *Groundwork,* 230 [4:429]. Distinguished scholars have interpreted these formulas in different ways, many of which I review in my essay "Kantian Normative Ethics," in David Copp, ed., *The Oxford Handbook of Ethical Theory* (Oxford: Oxford University Press, 2006), 480–514.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

sion procedures, I have suggested that a Kantian framework for assessing moral rules (or mid-level principles) might be constructed from Kant's idea of the perspective of a lawmaking member of a community of rational, autonomous, and mutually respectful persons (what Kant calls a *Reich der Zwecke* or "kingdom of ends").[21]

In section II of the *Groundwork for the Metaphysics of Morals,* Kant's attempt to articulate a basic moral standard unfolds in a progressive argument through the earlier formulas (of universal law and humanity as an end) to the idea of rational persons with autonomy legislating universal laws without basing them on contingent personal desires. Loosely paraphrased, the argument proceeds from "Conform to universal law" to "Respect and place the highest value on humanity in each person," followed by an explanation that this requires, among other things, respecting each person as a rational autonomous legislator of universal laws. This idea, in turn, is further developed in a formula of the Categorical Imperative that Kant suggests is his most complete characterization of moral requirements, the principle that says that we should conform to the laws of a possible *Reich der Zwecke,* or ideal commonwealth in which all rational autonomous persons, "abstracting from personal differences," legislate the moral laws that unite and bind them.[22] Although Kant still preferred his famous formula of universal law as the best guide for judgment in particular cases, he obviously regarded this later principle as an inspiring conception of the moral point of view, "closer to intuition" and "feeling."[23]

The main elements of Kant's idea of moral legislation are the following. The legislators are rational, seeking to legislate a consistent and coherent set of "laws" (moral principles), making no law without sufficient reason or in ignorance of relevant facts. They are committed to finding and respecting laws that apply to all and respect human dignity in all, and so, in legislating, they must rely on reasons that any moral agent can acknowledge and take into account apart from special interests. Their "autonomy" also implies freedom from various distorting and morally irrelevant influences and a genuine acknowledgment that the rational standard

[21] See note 1 for references to my fuller discussions of this Kantian perspective. According to David Falk in conversation, the frequent translation of *Reich der Zwecke* as "*kingdom* of ends" is due to H. J. Paton's insistence on using the term in his translation of Kant's *Groundwork* despite Falk's warning that the translation is misleading. H. J. Paton's translation is *Groundwork of the Metaphysic of Morals* (New York: Harper and Row, 1964). A *Reich* (literally) is a state or commonwealth, a governing legal system that may or may not involve a king or monarch. There is a possible analogy with biblical talk of a "kingdom of God" that may have influenced Kant but is not essential to his idea. When Kant describes the "kingdom of ends" (*Reich der Zwecke*) in the *Groundwork,* it has a "head" with features only possible for a divine being, but the head has no function apparently but to endorse the same rational principles as every other member. The primary difference is that the "head," who is not conceived as a political sovereign with independent legislative authority, lacks sensuous temptations to act contrary to reason and so is not "subject to" or "bound by" the rational principles. That is, they are not moral "imperatives" for the divine being.

[22] Kant, *Groundwork for the Metaphysics of Morals,* 231–37 [4:431–37].

[23] Ibid., 237 [4:436–37].

behind particular moral reasons is not merely rational prudence or efficiency in the pursuit of one's own personal ends. Obviously, and for good reason, certain moral values, assumed to be fundamental, are built into this conception of moral legislation. The aim is not to give a bootstrapping argument for basic moral values but rather to construct from these an intuitively appealing and inspiring conception of the moral point of view that should frame our thinking about more specific moral requirements.

A more modest application of Kant's idea, akin to Rawls's initial theory of justice, would make use of the model of "legislation" from an ideal perspective adapted for a more specific range of practical problems. For example, we might use this approach in addressing the question whether the prohibition of torture is open to exception. Given certain basic moral values, the presumption against torture (as well as against murder, rape, fraud, and manipulation) is rather obvious. The difficult questions are about cases in which moral values conflict. Granting that some particular hard cases may have no resolution, we can still ask whether it is morally important to maintain shared public principles about such matters, independently of law, and what exceptions, if any, these principles should allow. The broadly Kantian idea is that it may be helpful to treat such questions in two stages—articulating the relevant standpoint for joint deliberation (in effect, the rules of debate) and then reviewing arguments for proposed principles.

From the Kantian perspective, the most fundamental values that are to guide and constrain deliberation about rules are drawn from the idea of human dignity, or *humanity as an end in itself*. In Kantian theory, as I understand it, saying that something is valuable is equivalent to saying how we ought to regard and treat the thing that is said to have value. We can interpret the idea of human dignity, then, as a cluster of prescriptions about how to regard and treat human beings. In an earlier essay, I summarized these prescriptions as follows:[24]

> In sum, they are, *negatively*, not to regard persons as having a merely commensurable value and not to violate principles that rational autonomous persons would endorse, all things considered; and, *positively*, to exercise, preserve, develop, and honor rational nature in all persons, to prioritize the needs for rational autonomous living over more trivial concerns, and to count the happiness of others (or, strictly, their permissible ends) as significant in our deliberations. The proposed Kantian perspective regards the second negative prescription as an inflexible constraint on action: *never* violate the principles that, in one's best judgment, rational autonomous deliberators would

[24] Hill, "Assessing Moral Rules: Utilitarian and Kantian Perspectives," 168–69; emphasis in original.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

endorse, if guided and constrained by the idea of human dignity and mindful of the complexities of the real world for which the principles are intended. The first negative prescription ("*never* regard persons has having a merely commensurable value") and the remaining positive prescriptions are also non-optional demands, but they directly concern our *attitudes* when we deliberate and act.[25]

Obviously much more needs to be said, but it may be helpful to conclude my sketch with a few general remarks about a significant difference between the Kantian perspective and some theories that are in other respects similar. For example, the Kantian framework is meant to reflect what deliberation from a moral point of view about the issue at hand would be like, not necessarily what rational self-interested agents would choose in a state of nature or what "mutually disinterested" rational agents would choose under a "veil of ignorance."[26] Because those who deliberate from the broadly Kantian perspective are directly moved and constrained by basic moral considerations, the Kantian perspective, as construed here, resembles Locke's social contract theory more than

[25] Of course, they will rule out certain *acts* as immoral and show that others are morally required, but they do so by demanding that our deliberations, choices, and intentions are compatible with the prescribed attitudes. It is not as though one could muster up a motivationally inert and non-reason-giving "inner state of mind," such as a mere wish or "pure thought," and then anything one does will be permitted.

[26] As Fred Miller has noted in discussion, theories that aim to derive specific principles from a "self-interested" or any other deliberative point of view will not be able to reach any substantive conclusions unless they explain the interests or values that motivate the parties taking the deliberative (or contractual) perspective. "Self-interest" can be conceived in different ways, for example, as having the most pleasure possible, as realizing one's most important personal ends, or as Aristotelian *eudaimonia*. In Hobbes's state of nature, those who make the social contract are conceived as moved by their various desires, especially the predominant desire to survive, and their desires are always for some "good" for themselves. In Rawls's "original position," the parties are exclusively motivated to secure, within the constraints of the choice situation, the most favorable mix of "primary goods" (e.g., wealth, income, opportunities, powers, and self-respect) for themselves. They are mutually disinterested for purposes of deliberation under a veil of ignorance that bars them from knowing their special talents and circumstances, but they "represent" themselves as actual persons in the real world who may have deep religious convictions, family ties, and even altruistic dispositions. Both Rawls and Hobbes (on my reading) conceive of the parties in the relevant deliberative position (Rawls's original position and Hobbes's state of nature) as *not* motivated by explicitly moral concerns. The choice situations are deliberately defined to ensure that *we* can draw moral conclusions from arguments that the parties in those choice situations would endorse certain principles, but *the parties themselves* are not engaged in moral reflection. There are some theoretical advantages in conceiving of the parties' motivations as nonmoral, but, following Kant, my broadly Kantian deliberative perspective nevertheless conceives of the "legislators" of specific moral principles as motivated in part by stipulated basic moral values (for example, humanity as an end in itself). This approach, of course, invites questions about how the stipulated values can be defended and whether, given that moral values are stipulated, arguments for specific principles are circular or redundant. My brief discussion here suggests Kantian responses, but these questions obviously require more attention.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

Hobbes's.[27] In this respect, the Kantian perspective also appears to differ from T. M. Scanlon's contractualism, for Scanlon stipulates that in determining what principles can reasonably be rejected we must restrict ourselves to comparing the weight of reasons that different agents have *from their perspective*.[28] The reasons for endorsing or rejecting a principle that must be weighed are not *moral* reasons, even though the outcome of Scanlon's contractualist reasoning is supposed to be judgments about what is (or is not) morally wrong. My broadly Kantian approach, then, is just one of several ways of thinking about morality that attempt, first, to define an appropriate point of view for deliberation and debate and then, second, to use this perspective to derive specific principles about how to deal with various recurring practical problems. My approach differs from various contract-oriented theories, however, in explicitly calling for reliance on basic moral values in deliberating about specific principles.

IV. Strategies for the Defense of the Deliberative Perspective

To sketch Kant's strategies, we must review three main points: a background of critical arguments and then two different lines of defense for Kant's own perspective. The background (sketched in subsection A below) consists of objections to previous moral theories. Then the first positive strategy (subsection B) is an attempt to show "analytically" that the basic values in Kant's deliberative perspective are presupposed in concepts essential to our ordinary understanding of morality. The second positive

[27] John Locke's political philosophy employs the idea of reasonable agreements or contracts that were, or would be, made by persons who take for granted certain fundamental moral norms (the "laws of nature") as requirements of reason. Locke then tries to justify various specific principles regarding government, revolution, etc., by arguing that these principles would be endorsed by contracting parties who have personal interests but are also committed to these laws of nature (as moral guides and constraints). Thomas Hobbes, by contrast, tries to justify principles regarding government, law, etc., by arguing that rational persons in a state of nature would have sufficient prudential grounds, apart from any *independent* moral commitments, to accept those principles to be enforced by a sovereign power. Interpretations differ, but on my reading Locke, like Kant, is explicitly arguing from basic general moral values to more specific moral principles. Hobbes is apparently more concerned to avoid potentially controversial assumptions about moral values and instead to justify specific principles as rationally prudent to accept and follow if one is assured that others will. Among those principles are Hobbes's "other laws of nature," many of which resemble familiar moral norms, but Hobbes apparently regards these as rationally binding only because following them serves to enhance one's power to survive and satisfy one's desires. See John Locke, *Second Treatise of Government,* in *Classics of Moral and Political Theory,* ed. Michael L. Morgan (Indianapolis, IN: Hackett Publishing Co.), esp. chaps. 1–9; and Thomas Hobbes, *Leviathan,* part I, esp. chaps. XIV and XV. We should note, however, that I have been comparing a Kantian *moral* perspective with the *political* theories of Locke and Hobbes. A comparison of the latter with Kant's political and legal theory is complicated because the relation between the Categorical Imperative and Kant's principles regarding law is unclear and controversial.

[28] Scanlon, *What We Owe to Each Other*, esp. chaps. 2, 4, and 5.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

strategy (subsection C) is simply to appeal to each reader's moral consciousness, which Kant thought was a virtually inescapable recognition of being under moral obligation (as Kant conceived this).[29] Apart from the details of Kant's own exposition, arguably these general background arguments and strategies are potentially relevant in assessing contemporary moral theories, including the broadly Kantian perspective sketched here.

A. The background: Critique of previous moral theories

The background consists of negative arguments to reject other (non-Kantian) views about the nature and source of morality. These arguments are most explicit in Kant's *Groundwork for the Metaphysics of Morals* and *Critique of Practical Reason*.[30] Kant categorizes all previous moral theories into four types, which we may label "ethical egoism," "moral sense theories," "divine command theories," and "perfectionism." Having argued that moral concepts presuppose that "autonomy of the will" is the fundamental source of morality, Kant charges that all the previous moral theories are mistaken because they fail to acknowledge this fact. Here, I take it, Kant is engaging in what is now considered "metaethics," for he maintains that his "analytical" argument in *Groundwork* II shows that the theories in question fail to understand common moral concepts and, further, that they leave us without any resources to argue (as Kant attempts in *Groundwork* III) for the rational authority of basic moral standards. Kant alleges that each type of moral theory is open to special objections, but the common error of each theory is to suppose that morality is something other than following principles that moral agents with practical reason and autonomy would give (or "legislate") to themselves as constraining their social relations and personal choices. We do not need to accept the full Kantian idea of autonomy to appreciate his critique of other moral theories.

Roughly, the main negative points, reworked and modified for contemporary discussions, are these: (1) Some moral theories attempt to derive the content and rational authority of moral principles from the idea that, fundamentally, each person should (or has sufficient reason to)

[29] Those who were looking for a logical demonstration or empirical confirmation of Kant's basic value claims are likely to find this appeal to their moral consciousness initially disappointing and perhaps not even worth calling a "defense." It also falls short of Kant's ambitious project in *Groundwork* III to "establish" that anyone who acts for reasons does so "under the idea of freedom" and is therefore committed to the Categorical Imperative as a supreme standard. However, given a background of arguments against other moral theories, the attempt to get us to see, in our own reflections, that we deeply and persistently acknowledge certain basic moral values as rational might suffice, for practical purposes, in response to certain kinds of theory-driven doubts about whether acting morally makes sense.

[30] See Kant, *Groundwork for the Metaphysics of Morals,* 242–45 [4:441–44], and Kant, *Critique of Practical Reason,* 32–36 [5:35–41].

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

do whatever most effectively promotes his or her interests. This approach is inadequate, Kant argues, because it is a constitutive conception of *moral* requirements that they prescribe what we must do independently of whether they advance our own interests. Thus, even if only the pursuit of self-interest is rational, the requirements that would be prescribed by this method would not be *moral* requirements. Kant classifies egoistic theories as "empirical," apparently because their advocates typically base them on psychological egoism[31] and because they make specific moral requirements vary widely with each person's empirical circumstances. Kant's main critical point, however, is that moral requirements cannot be derived from egoistic principles of any kind, rational or otherwise.

Further, (2) Kant also rejects moral sense theories, presumably in part for reasons that comparable theories tend to be rejected now. Such theories attempt to derive moral conclusions from patterns of moral feelings that human beings typically have in recurrent situations—for example, impartial spectators' feelings of disapproval of cruelty. The core objection is that judgments about what we ought to do are *normative judgments* about what we have *good reasons* to do, and these do not follow (without further premises) from facts about human sentiments, no matter how universal these may be. The point for contemporary discussions would be that empirical psychology alone cannot establish moral claims about what we ought to do, even though it can undermine ill-informed particular moral judgments and cast doubt on our capacity to live by certain alleged moral standards. Assuming (with Kant and others) that moral claims purport to give us good reasons to act, empirical studies by themselves cannot determine which moral claims, if any, are true or valid.

Turning to theories that Kant classifies as nonempirical or rational theories, (3) he rejects theories that attempt to derive moral requirements from theology on the ground that either their derivations are circular or the derivative commands are not moral. In other words, either the divine source is *defined* as giving only rational moral commands or else its commands are nonmoral because the authority comes from divine power, not reason. This objection also poses a challenge for any contemporary moral theory that rests the ultimate reasons for being moral on actual commands, secular or ecclesiastical. Similar considerations might apply to the assumption that the "force" of tradition, custom, or common opinion (alone) is the basis of moral requirements.

Finally, (4) Kant rejects forms of perfectionism that base the rational authority of moral requirements on their alleged promotion of the metaphysical value of "perfection," an idea prominent in the work of Leibniz and Christian Wolff. This idea, Kant suggests, is too indeterminate to be

[31] Psychological egoism, as I understand it, is the view that, as a matter of fact, every person is always ultimately motivated by the desire to advance his or her own interests. It can take different forms, depending on how "one's own interests" is interpreted and how deep or near the surface self-interested desires are taken to be.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

action-guiding unless we interpret "perfection" in a morally loaded way that would make derivations of moral requirements from it circular or question-begging. Again, details aside, we can draw a message relevant in moral theory today. Metaphysical properties, whether labeled "perfection" or "intrinsic value" or whatever, are supposed to be features of *the world as it is,* not necessarily *as it ought to be.* What we have reason to do cannot be determined by such facts alone. Although we are not concerned with the sort of perfectionism that Kant explicitly rejected, his objection may be relevant to current theories that rest too easy with "intuitions" about which facts are practical "reasons" (or "considerations that favor" choices), in the absence of any general account of what constitutes reasons of various kinds.

Returning now to my review of Kantian strategies for defending alleged basic moral standards, the first main point (which provides background for more positive arguments) has been that certain other strategies are *inadequate for the defense* of Kantian basic standards and, equally important, that those strategies *cannot by themselves refute* Kantian proposals about what the rational moral standards are. This is a strong negative metaethical claim that, perhaps for good reasons, most contemporary Kantians do not wish to debate at length, but it is a background idea that, if correct, makes it easier to move on to positive considerations.

B. Argument that the perspective is presupposed in common moral judgments

Kant's main attempts to defend his conception of the basic moral standard are presented in the first two chapters of the *Groundwork,* where he argues that concepts deeply embedded in common morality presuppose the supreme principle expressed in his versions of the Categorical Imperative. His aim in these chapters is not to "establish" his basic principle or standard as rational but only to identify formulations of that standard that express its essential features and can serve to guide moral deliberation. If successful, these arguments taken by themselves would provide an important but limited "defense" of Kant's basic deliberative perspective as reflecting the indispensable core of our common morality, rather than mere custom or personal preference.

Kant's arguments do not start from a survey of moral opinions about particular matters, such as lying and suicide, but from relatively formal ideas about duty and a good will. An essential feature of these ideas, supposedly pervasive in ordinary moral thinking, is that being good and doing what one ought is, among other things, a matter of respecting constraints that we have sufficient reason to follow and not just because doing so advances our interests. This could be so, Kant argues, only if there is a basic rational standard, distinct from instrumental and prudential rationality, which can support more particular moral judgments. More-

over, Kant argues that the content of the basic standard(s) is expressible in different ways, which are supposed to be fundamentally the same. Most abstractly expressed, the standard is that we must (rationally) choose in a way that respects "universal laws." As eventually interpreted, the standard is to respect the more specific principles of conduct that would be "legislated" by rational agents who have "autonomy of the will" and respect humanity in each person. The form of Kant's argument is conditional, and so is his conclusion. That is, essential concepts of common morality *presuppose* his basic deliberative perspective as the standard from which genuine moral obligations derive. This defense is distinct from his later attempt, in *Groundwork* III, to show that common morality is not mistaken in taking its requirements to be *rational* to follow.

Let us set aside details and doubts about Kant's actual argument (and my brief summary) and focus instead on the kind of positive strategy Kant uses to defend his attempt to articulate a basic moral standard. Kant's strategy is similar in some respects to Rawls's method of reflective equilibrium, but there are also important differences.[32] Some differences in their methods stem from differences in their ambitions and hopes for moral theory, but others reflect Rawls's greater respect for the empirical realities which influence and provide the context for moral reasoning. These differences understandably make Rawls's method more acceptable to contemporary philosophers, but the ways in which Rawls's method remains similar to Kant's are also instructive.

To set the background very briefly, Rawls draws from widely accepted, relatively uncontroversial assumptions about basic moral and political values to develop (or "construct"?) an appropriate idealized point of view ("the original position") for assessing competing conceptions of justice.[33] He proposes two general principles of justice and then argues that these would be preferred to certain alternatives from the perspective of the original position. Then Rawls argues that applications of the principles, in several stages, more or less match the "considered judgments of competent judges" in due reflection. When arguments from the appropriate deliberative standpoint ("the original position") *as this was initially defined* lead to conflicts with our "considered judgments," we need to reflect further on the definition and arguments for the original position, its application, and our own pretheoretical considered judgments. The task of seeking *reflective equilibrium* is to revise the arguments, refine the original position, or abandon our initial moral judgments, until we can find no more acceptable way to bring our various convictions into a coherent whole. Like Kant, Rawls argues from what he takes to be deep and pervasive ordinary moral ideas to support an account of an appropriate

[32] John Rawls, *A Theory of Justice, Revised Edition* (Cambridge, MA: Harvard University Press, 1999), 18–19, 42–45, and 507–8.
[33] Ibid., esp. part I, 10–19.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

deliberative perspective for assessing more specific principles, and then he illustrates the application of the deliberative perspective by applying it to several practical problems.[34]

Significant differences, of course, should be noted. Most obviously, Kant's ambition was to present a comprehensive moral standard, whereas Rawls aims to construct a deliberative perspective for a more limited project. That is, even in *A Theory of Justice,* Rawls constructs his "original position" primarily to assess competing, historically prominent conceptions of justice as they apply to the basic structure of well-ordered, cooperative, closed societies—in particular, with regard to how the basic structure affects the distribution of primary social goods. Later, in *Political Liberalism,* he restricts his project further to "political conceptions" of justice that reflect background assumptions common in Western liberal democracies.[35] A significant feature of Rawls's arguments for his deliberative procedure for assessing principles is that, unlike Kant, Rawls does not try to make his case by drawing out the implications of a few, relatively formal and allegedly essential features of common morality (such as the idea of duty and the value of a good will). Rawls rightly acknowledges more places where empirical facts are relevant to his arguments, and he is obviously more aware of the possibility that revisions may be necessary.

Important similarities remain, however—and for good reasons. Both Kant and Rawls see moral theory as a project with a structure that distinguishes basic deliberative procedures from derivative principles, and both distinguish arguments *for* the deliberative procedure from arguments *from* the procedure to derivative principles. Although both rely on assumptions about common moral opinions, neither treats these opinions as self-evident or infallible. Despite his repeated warnings about not basing moral theory on examples, Kant (like Rawls) draws out the (supposed) implications of his principles as some sort of confirmation of them.[36] Neither Kant nor Rawls addresses his theory to radical moral skeptics or pretends to have arguments sufficient to show that living a moral life is necessarily in one's interest. Finally, both Kant and Rawls seem to realize that, in the end, defense of their projects and results

[34] Rawls takes up various practical problems in part II of *A Theory of Justice,* for example, toleration of the intolerant (190–94), governmental institutions to promote just distribution of wealth and income (242–51), justice between generations (251–58), and civil disobedience (319–31).

[35] John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993).

[36] Kant says: "[M]y claims about this central question [the nature and justification of the supreme principle of morality] would be greatly clarified by seeing the application of that supreme principle to the whole system, and they would be *strongly confirmed* by the adequacy the principle would manifest throughout." Kant, *Groundwork,* 193–94 [4:492]; emphasis added. He nevertheless strongly condemns any attempt to *derive* fundamental moral principles from examples. See, for example, *Groundwork,* 208–14 [4:406–12]. *The Metaphysics of Morals* contains his most systematic attempt to present and support a system of intermediate-level principles.

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

depends on a different sort of appeal to each reader—the subject to which I turn next.

C. Appeal to each rational agent's awareness or acceptance of the basic standards

A final strategy for defending an account of basic moral deliberative standards is suggested by Kant's attempts to address the question that he so dramatically poses at the end of *Groundwork* II. That is, even if an account is demonstratively reflective of common moral ideas, is it really *rationally necessary* to accept it? Might it not be that morality's claim to be rational is just an illusion? Kant tried in *Groundwork* III to meet the challenge by arguing that whenever we act for reasons, we cannot help but see ourselves as free in a sense that ultimately commits us to the Categorical Imperative.[37] To meet the objection that he may have argued in a circle, he adds that in the theoretical use of reason (for example, in scientific investigations) we must take ourselves to be following reason independently of causal determination, and this again commits us to the Categorical Imperative.[38]

Later, in his *Critique of Practical Reason,* Kant reverses his argument, apparently conceding the inadequacy of the early arguments. He then proposes that consciousness of the moral law is "a fact of reason," adequate for practical purposes.[39] Scholars debate what he meant, but apparently at least part of the idea is that for us, human beings who take ourselves to be rational moral agents, no further argument is needed because, try as we will to deny and escape moral requirements, we cannot. That is, although self-deception, special pleading, and moral disagreements are all too common, in serious honest reflection we cannot consistently regard ourselves as rationally entitled systematically to ignore the claims of others as not (rationally) binding on us. The rational authority that supports such claims, Kant suggests, is inevitably recognized as the authority of our own reason, however much we may wish to deny it. This is not presented as a demonstration, or a claim to perceive self-evident rational truths, or an appeal to moral feelings. As I understand Kant's point, it is an appeal to individuals, in their own first-person reflections, to acknowledge that they are under moral obligations or at least cannot help but conceive of themselves in this way. Arguably, despite Kant's own views, such reflections should not be used to justify assuming that others are as committed to morality as we are—assuming this, for example, when we are inclined to justify harsh moralistic treatment of criminals without regard to their actual background and attitudes. Never-

[37] Kant, *Groundwork for the Metaphysics of Morals,* 245–48 [4:444–48].
[38] Ibid., 248–57 [4:448–58].
[39] Kant, *Critique of Practical Reason,* 28–64 [5:30–50].

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

theless, Kant apparently thought his appeal to each person's moral consciousness a relevant kind of response to certain kinds of doubts about morality, given the fact (argued earlier) that contrary claims about rational choice could not be proved by science, theology, metaphysics, or social practice.

Kant's expectations for the practical success of this appeal to each person's moral consciousness were higher, I suspect, than the expectations most of us share. This is especially so if the person's moral consciousness is supposed to acknowledge Kant's particular statements of the moral law and associated ideas about freedom of will. But at least when our aim and context is practical, rather than an attempt to establish a universal truth of metaphysics or abstract moral theory, we may find that the best we can do, and all that is needed, is to appeal to the commitments of each rational moral agent. If agents profess to deliberate and debate in terms of shared reasons, reasons about which others outside their circle should care, then they may be able to see, on due reflection, that their moral and reason-based commitments exceed their wishes to escape moral obligation. And perhaps a more modest, reasonable contemporary account of moral obligations (compared to some ideas in Kant's actual texts) can facilitate this process, and arguably it can be a (practically) rational process. Rawls's appeals to each reader to consider, revise, and endorse his theory of justice is similar and, in my view, remains worthy of serious consideration.

V. Conclusion

In this wide-ranging essay, I began by distinguishing different questions that are often confused in discussions of moral relativism and objectivity, and then I attempted to sketch the responses that a broadly Kantian theory might give to two of these questions. The first of these questions was: How (if at all) can we derive, justify, or support specific moral principles and judgments from more fundamental moral standards and values? In response, I briefly sketched a broadly Kantian deliberative perspective designed for this purpose, drawing from Kant's later formulations of the Categorical Imperative. The second question was: How (if at all) can fundamental standards such as those found in the Kantian perspective be defended? In response, I briefly described two of Kant's strategies for defending his fundamental standards and the important background of arguments against previous moral theories.

The Kantian deliberative perspective, like rule-utilitarianism, sets us the task of finding or constructing a set of specific moral principles (or "rules") that satisfy basic values and constraints but also take into account relevant facts about the human condition and the circumstances in which the principles are to apply. Unlike rule-utilitarianism, however, the Kantian perspective essentially includes values associated with human dignity

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press

that shape Kantian deliberations about the set of specific principles to be adopted. Kant's strategies for defending the basic values inherent in the deliberative perspective are, first, to try to show that these values are presupposed in common moral thought and practice and, second, to appeal directly to the moral consciousness of the would-be skeptic. In neither strategy does Kant argue for his standards from (allegedly) more self-evident premises or appeal to intuitive "perception" of an independent moral order. Instead, he analyzes the presuppositions of common moral concepts and appeals to each person to acknowledge his or her own deep and persistent practical commitment to the basic standards as rationally authoritative. How far such arguments can succeed in the world today is an open question. Kant's own confidence in them was a kind of morally based faith not refutable by empirical studies, but it is hard to share that confidence in today's world.

Philosophy, University of North Carolina at Chapel Hill

https://doi.org/10.1017/S0265052508080084 Published online by Cambridge University Press