Synthese (2021) 198 (Suppl 13):S3155–S3170
https://doi.org/10.1007/s11229-018-01977-x

S.I.: THE CURRENT RELEVANCE OF KANT'S METHOD IN PHILOSOPHY

Methodological reflections on Kant's ethical theory

Robert Audi[1]

Received: 10 December 2017 / Accepted: 5 October 2018 / Published online: 1 December 2018

Abstract

In an important passage in Kant's *Groundwork,* he says: "[W]e cannot do morality a worse service than by seeking to derive it from examples. Every example of it presented to me must first itself be judged by moral principles in order to see if it is fit to serve as an original example—that is, as a model: it can in no way supply the prime source for the concept of morality" (4: p. 408). This is an important methodological pronouncement, and it appears to commit Kant to what might be called a "top-down" procedure for constructing an ethical theory—or at least for defending substantive moral principles. A contrasting method we might call "bottom-up" would attribute to what are commonly called *intuitions*, especially those concerning concrete cases, a basic epistemological role in such a theoretical normative project. This paper undertakes, first, to clarify both kinds of procedure and to sketch a philosophical methodology that can do justice to certain merits of each procedure; second, to explore, drawing on a methodological analysis the paper will outline, Kant's actual *operative* method in much of his ethical writing, particularly but not exclusively the *Groundwork*; and third, to appraise some aspects of Kant's actual methods of theory-building as it is seen in his development of his ethical framework. The concluding reflections will show that Kant's overall achievement in moral philosophy does not depend on certain of his metaphilosophical views. The paper will also indicate some directions of moral inquiry that may be promising for both Kantian and other approaches in moral philosophy.

Keywords Bottom-up methodology · Good will · Intuition · Rigorism · Top-down methodology

✉ Robert Audi
audi.1@nd.edu

[1] Notre Dame, USA

Kantian ethics, like virtue ethics and utilitarianism, is now even richer than the position of its founder. The literature on it, both as focused on Kant's writings and as defending or extending him, is immense. This paper does not survey that literature. Its aim is restricted: to clarify Kant's methodology in moral philosophy and, so far as possible, to advance our understanding of both what the methodology is and how it may constrain his normative ethics. To do this, we must briefly describe two contrasting methods of ethical theory-building and explore them in Kant's ethical writings. We must also consider both his own methodological statements and a quite different element in his work: his actual procedure in formulating, illustrating, and defending some his views.

1 The top-down versus the bottom-up in philosophical methodology

I take methodology to be the theory of method and will describe two major kinds of philosophical method useful in understanding Kant. Methods embody procedures but are systematic in a way procedures need not be, and I assume that an overall philosophical method will concern ways to achieve at least these interrelated aims: formulation, discovery, explanation—including systematization—and defense. With these aims in mind, I take the main methodological elements in top-down philosophical approaches to reflect a quadruple preference: for generality over particularity, for the a priori over the empirical, for the necessary over the contingent, and for the axiomatic over the derivative. Bottom-up methods may in principle reverse all of these preferences, but perhaps no philosophical method is entirely either top-down or bottom-up. In any case, a major element in many bottom-up methods is giving some priority to certain kinds of examples, particularly the kind important in applied ethics.

With these points in view, consider the passage in the *Groundwork* in which Kant says, "[W]e cannot do morality a worse service than by seeking to derive it from examples…" (4: p. 408).[1] The passage apparently presupposes that we may have (and presumably even know) *principles* prior to seeing an example that instantiates them. We can, for instance, see that lying in general is wrong before seeing the wrongness of any particular lie. In brief, *generality is prior to particularity.* This priority is *epistemic,* not temporal. One cannot, e.g., know that a specific action, such as a lie, is wrong except on the basis of knowledge that it is of a *type* that is generally wrong.

Generality does not entail necessity, but Kant leaves no doubt that moral "laws" are necessary (and he apparently includes as laws moral principles in general, at least assuming their derivability from the Categorical Imperative). This requirement is stated often in the *Groundwork* among other works: its preface says, "[A] law has to carry with it absolute necessity if it is to be valid morally—that is, as a ground of obligation" (4: p. 389).[2] His example here is "Thou shalt not lie." This and many other examples indicate that moral laws may be quite ordinary directives.

[1] Here and elsewhere in referring to this book references will be to sections numbered as in the Prussian Ed. and, unless otherwise specified the translation is H. J. Paton's *The Moral Law: Kant's Groundwork of the Metaphysics of Morals* (London: Hutchinson & Co., 1948). For detailed discussion of Kant's use of examples, see Robert B. Louden, "Making the Law Visible: The Role of Example in Kant's Ethics," in Louden's *Kant's Human Being* (Oxford: OUP, 2011)," pp. 91–104.

[2] In 4 [the *Groundwork*]: p. 421 Kant says conditionally that "all imperatives of duty can be derived from this" (the universal law formulation in that section); but he surely holds this. Cf. 4: p. 429 where, having

Generality also does not entail apriority. But Kant considers apriority to be, like generality and necessity, a requirement for moral laws (and this requirement does not depend on whether apriority is entailed by necessity). The same passage says, "[T]he ground of obligation must be looked for, not in the nature of man nor in the circumstances … but solely a priori in the concepts of pure reason" (4: p. 389). Later Kant says, "Since moral laws *have to hold* for every rational being as such, we ought rather [than derive them from the particular constitution of human reason] to derive our principles from the general concept of a rational being as such…" (4: p. 412, italics added). Moral laws, then, are a kind of conceptual truth that (though it need not be analytic) is both a priori and necessary.

Kant appears to leave open the genetic question of how we acquire the concept of a rational being, but the central role of concepts in theory-building is affirmed earlier, in the passage just quoted: "All moral concepts have their seat and origin in reason completely a priori, and indeed in the most ordinary human reason just as much as in the speculative; they cannot be abstracted from any empirical, and therefore merely contingent, knowledge" (4: p. 411). Methodologically, Kant seems here to affirm a conceptual priority of reason over experience of particular cases to which the relevant concepts apply. This view does not entail a Platonic ontology, however, and I leave open how Kant might connect the a priori with the distinction between the abstract and the concrete.

Such passages make it natural to describe Kant's method as *top-down.* Metaphor is often risky, but this one is useful if taken to indicate adherence to the quadruple priority just sketched: of the general to the particular, the necessary to the contingent, the a priori to the empirical, and the axiomatic to the derivative. This is a priority he considers conceptual as well as epistemic. These elements may all be implicit in a passage in the Second Critique:

> The method [of determining our duty] therefore takes the following course. The first step is to make judging according to moral laws a natural occupation which accompanies our own free actions as well as our observations of those of others, and to make it, as it were, a habit. We must sharpen these judgments by *first asking whether the action is objectively in accordance with the moral law*, and if so, which one… The second point… is whether the action is done (subjectively) for the sake of the moral law… reason, with its faculty of *determining according to* a priori *principles what ought to occur*, can find satisfaction only in such an order of things.[3]

A question that now arises is what constitutes a bottom-up method that instructively contrasts with Kant's.

Suppose we take the verticality metaphor seriously and think of generalizations and "general" concepts as, methodologically speaking, at the top level in our theorizing and concrete examples as at the bottom level—as where, in ordinary life, a child tells a lie and is told, e.g., 'That's lying, and it's wrong!' Historically, virtue ethics seems

Footnote 2 continued
described rational nature as the ground of the imperative (as the humanity formula), he says that from it "it must be possible to derive all the laws of the will."

[3] See Lewis White Beck, trans., *Critique of Practical Reason* (Indianapolis: Bobbs-Merrill, 1956), 5: pp. 159–160 (pp. 163–164). Mary Gregor's translation is consistent with this).

the ethical approach in which bottom-up methodology is most prominent: role models, mainly by example, teach morality through illustrating, by deeds perhaps more than with words, what is morally required (or best); and, from observing such particulars, learners acquire general moral knowledge of relevantly similar situations. Or, as just indicated, someone does a wrong or fulfills an obligation, and, from discerningly viewing such particulars, observers learn generalizations capturing a relevant connection—say, between promising to A (an action) and A-ing. Intuitionism in certain cases also seems bottom-up, perhaps giving to intuitions—not in the Kantian sense but in the current sense of non-inferential cognitions—about concrete cases the kind of role played by judgments of the *phronimos*—the person of practical wisdom.[4]

This bottom-up picture of the origin of moral knowledge is certainly inapplicable to the pictures most often painted by Kant. He says, e.g., "Precepts, not the example of others, should be the ground of our actions,"[5] though he grants that "On the whole, examples are desirable" (p. 111). But, as O'Neill put it, "On Kant's view actual cases of moral deliberation do not use examples at all… to decide what to do we are required to test the principle on which we propose to act according to the Categorical Imperative."[6] Precepts need not be presented apart from illustrations, but even illustrated general principles have epistemic priority over their illustrations.[7]

Might there be a bottom-up method in ethics quite different from the kind Kant rejects and representative of the leading kind in contemporary ethics? Here I turn to ethical intuitionism and begin with W. D. Ross, who took much from Aristotle but also represented a bottom-up method as in *one* way like a top-down method as we find it in Kant: namely, in having an a priori element. Ross illustrated it with learning from combining pairs of matches as a way of coming to see that $2+2=4$. For Ross, C. D. Broad, and others, ethical learning might go partly as follows. Imagine that one somehow grasps (as might be possible given observations of a role model) an individual ground as requirting a kind of act, say an action of promising to A as demanding A-ing. This grasp, as manifested in a sense of duty to do the promised deed, is a (possibly non-formulated) recognition of a reason. From this point, one is in a position both to conceptualize promising and to see the truth of the generalization that promising yields a prima facie "duty" (roughly, a moral reason). Now if this is the *only* possible way to learn moral concepts and moral generalizations, we have a sharp contrast with Kant's view. But neither intuitionism in general nor (arguably) Ross's version, is necessarily committed to that narrow view.

In Ross's bottom-up, particularistic portrayal of at least some moral learning, intuitive induction goes beyond yielding a grasp of moral generality. It can also yield at

[4] Regarding conflicting obligations, Ross, e.g., approvingly cites Aristotle's dictum in NE 1109b23, "The decision rests with perception." See W. D. Ross, *The Right and the Good* (Oxford 1930), p. 32.

[5] See Louis Infeld, trans., *Lectures on Ethics* (London: Methuen, 1930), p. 111. Page references are to Lewis White Beck's edition (NY: Harper Torchbooks, 1963).

[6] See her *Constructions of Reason*: *Explorations of Kant's Practical Philosophy* (Cambridge, CUP, 1989), p. 166. She later quotes him as saying that judgment—which is of course crucial for (among other things) framing the right maxims "cannot be taught" (op. cit., p. 167).

[7] Some qualifications are needed, but not crucial here. I offer an account of role modeling applicable to Kantian moral psychology in relation to learning moral principles in "Role Modeling and Reasons: Developmental and Normative Grounds of Moral Virtues," in Noell Birondo and Stewart Braun, eds., *Virtue's Reasons: New Essays on Virtue, Character, and Reasons* (New York: Routledge, 2017), 126–44.

least some apprehension of the *necessity* of the grounding relation. Here Ross is like other rationalist moral philosophers. As Broad said, Ross believes "Reason needs to meet with concrete instances of fitting or unfitting intentions and emotions before it can rise, by Intuitive Induction, to the insight that *any* such intention or emotion would necessarily be fitting (or unfitting) in *any* such situation."[8] If the 'need' is conceptual and not just empirical, Kant rejects the claim. He says, e.g., that "reason commands what ought to be done even though no example of this can be found in experience…"[9] Kant can, however, allow that experience provides exemplars and thereby raw material for acquiring a priori knowledge: it can be a genetic but not epistemic basis for such knowledge.[10]

2 The epistemology of Kant's ethical method

We have now seen good reason to consider Kant a top-down theorist in a strong sense, but we have also seen that the contrast between top-down and bottom-up methods may be less stark than the metaphors suggest. One point is that both Kantian ethics and classical intuitionism are epistemologically rationalist. Limitations on the role of particularity, however, above all in moral learning, is a point of disagreement. Let me briefly elaborate.

An intuitionist who is a rationalist will maintain that general moral principles are a priori, but is free to hold that this point is independent of how we come to understand or know them. Intuitionists tend to give high priority to examples in moral learning: the exemplificational "bottom" level. But appealing to intuitive induction is not the only way to account for the role of examples. They might lead to generalization without the conceptually distinctive abstraction characteristic of intuitive induction. There, as in other cases, the apriority of what is accepted on the basis of experience is independent of genetic factors. How generalizations can be known is logically independent of how they can be arrived at, whether through experience, intuitive induction, or a priori reflection.

Rationalists tend to take a priori propositions to be necessary (even if some reject the converse), and here, too, bottom-up theorists may agree with top-down moral philosophers. The difference, as illustrated by the contrast between Ross's intuitionism and Kant's moral theory, is in the place of examples in both moral learning and ethical theory. Even the major ontological question of grounding is left open: both kinds of theory, in contrast with naturalistic empiricist views such as Mill's, may take the grounding relation between, say, promising to A and being obligated to A as a priori and necessary.[11]

[8] C. D. Broad, *Five Types of Ethical Theory* (London: Routledge and Kegan Paul, 1930), p. 282.

[9] See Kant's Introduction to *The Metaphysics of Morals,* in Mary J. Gregor, trans., *The Doctrine of Virtue: Part II of the Metaphysics of Morals* (New York: Harper and Row, 1964), p. 13.

[10] Kant makes it clear in many places that the sense in which a priori knowledge is independent of experience does not entail that no experience is required for understanding the a priori propositions in question. See, e.g., CpR B1-2.

[11] For Ross, the obligation is prima facie; for Kant, perfect obligations are often represented as always final, but, as I will shortly explain, that rigoristic position is not essential to his view.

Maintaining that moral principles are necessary and a priori leaves open their interrelations. Kant surely considered the Categorical Imperative his master principle, in the sense that it is like an axiom in being a basis for deducing all other (moral) principles. This leaves open whether he endorsed hierarchy among the others. There are various kinds of hierarchy. Here my concern is whether he was committed either to (1) a *hierarchical ordering* that lists all the major types of obligations by importance, in the sense that obligations of different kinds (represented, in his system, by different principles) are always treated uniformly, with *every* singular (roughly non-compound) obligation in a higher category, say that of promissory obligation, overriding *any* single one in any lower category, say beneficence; or (2) a *pairwise hierarchy* such that, for some (but not all) such conflicts, such invariant dominance holds. In case (2), which does not imply a hierarchy of *all* major obligation types, promissory obligations might be "above" those of beneficence but on a par with obligations of veracity. I do not see that Kant was committed to a full hierarchical ordering. But, unlike Ross, he appears to endorse a limited deontic hierarchy regarding perfect duties in relation to imperfect duties.

There is a related epistemological contrast between Kant's top-down method in ethics and the intuitionist method of Ross and many other philosophers, including many non-intuitionists. The difference concerns conflicts of prima facie obligation.[12] Consider the most commonly cited obligational conflicts: one makes a promise yet finds that keeping it conflicts with a serious failure in the duty of beneficence, or finds that the only way to prevent a serious wrong is to lie. Rossians and virtue ethicists hold that moral obligations are (at least as to weight) *incommensurable* and that, accordingly, such cases call for judgment in the light of the totality of relevant considerations. This view apparently implies that these cases are resolvable only in what seems a bottom-up fashion—bottom-*up* rather than only *at* the bottom level because one may, after the fact (say after resolving an obligational conflict), be able to frame a generalization that holds universally and this may properly influence one's final judgment on the conflict. Kant, on some readings, takes deontic conflicts to be resolvable on the basis of an antecedent generalization, hence in a top-down fashion. He says, for instance (concerning the temptation to lie to avoid evil), "[I]mperfect duties always succumb to perfect duties, just as several imperfect duties outweigh a single one; for example, the distress of another, were it even to be mortal, could not compel me to contract debts…"[13] and implicitly endorsing a pairwise hierarchy in which obligations of veracity override those of beneficence in conflict cases—that "an action of this kind [of an imperfect duty overriding a perfect one] must be considered bad in itself, and that the imperative of prohibition is therefore categorical" (4: p. 419).[14] The quoted statement follows shortly after he affirms that "it is impossible to settle *by an example,* and so empirically, whether there is any imperative of this kind…" (4:

[12] Granted, Kant says, "A *conflict of duties*… is inconceivable." See *The Metaphysics of Morals* (6: p. 223). But Kant quickly adds that two "grounds of obligation can conflict" (ibid.). That is all Rossians require, since it is only prima facie and not final duties that they take to conflict. Passages supporting the compatibility of Kant's view with Ross's appear on pp. 261, 273, and 296 in Peter Heath and J. B. Schneewind, eds., *Immanuel Kant: Lectures on Ethics* (Cambridge: CUP, 1997).

[13] See the Virgilantius notes, in Heath and Schneewind, op. cit., p. 296.

[14] Gregor has 'evil in itself', but the difference does not affect my point here.

p. 419). Similarly, in the *Lectures* he said, "The dictates of morality are absolute… Our free doing and refraining has an inner goodness… [which] endues man with an immediate, inner, absolute moral worth" (p. 5).

For intuitionists, a natural view here is that even if, apart from exercising intuitive judgment, we cannot know which (if either) of two incommensurable conflicting duties prevails, still, if, on the basis of intuitive discernment, we do know which prevails, then we can *in principle* formulate a generalization (a kind of maxim) that the case instantiates. There is, to be sure, a problem that Kantian ethics must address: "it is difficult to find any way of characterizing the proper description of the maxim without relying upon one's antecedent sense of how the test should come out."[15] For intuitionism, by contrast, the list of grounds of prima facie obligations is a constraint on framing the relevant generalizations (corresponding to Kantian maxims), even if there remains a measure of indeterminacy as to what obligation is final. On this conditional claim—taken epistemically—Kant might agree. Even if he believed that one never has more than one (non-disjunctive) obligation on a single occasion, he need not have held that we can always know what our actual maxim would be on a given occasion, or whether a maxim is (rationally) universalizable. If one assumes that, as might be thought implicit in his top-down, aprioristic method for determining particular moral obligations, he is a rigorist—allowing no exceptions to demands of perfect obligation—one would think him committed to holding that perfect obligations are always overriding. But this would overlook the point that *exceptions* to such obligations as promise-keeping might themselves be justified by rationally universalizable, if narrowly applicable, principles. Let me explain.

First, let us make the plausible and widely accepted assumptions (apparently open to Kant) that (1) obligatoriness is grounded in natural properties (such as killing) discernible in the situation calling for moral decision, and (2) moral judgment is properly responsive to those properties. Even for Rossians, but especially for theorists who do not posit an a priori ascertainable complete list of relevant grounds, it may not be clear what *all* the relevant grounds are; but, commonly, situations of moral decision exhibit enough grounds for obligation to enable rational determination of what to do. Kant might or might not see this grounding idea as harmonious with his ethics, but his view apparently differs from the intuitionist position in at least two respects. First, he would not give the intuitive, particularistic judgments of obligation *independent* epistemic weight. Second, he would tend to dispute the incommensurability of conflicting considerations—or at least to hold that resolving those conflicts is not a matter of intuitive singular judgment. Rather—thirdly—resolution requires determining a universalizable maxim. Section 3 addresses these points.

3 The methodological rationalism of Kant's ethical theory

Consider the rigoristic principle 'Suicide is always wrong', which Kant might be thought to hold. Could he not consider rationally universalizable, *for* nuclear missile

[15] See Thomas E. Hill, Jr., *Respect, Pluralism, and Justice: Kantian Perspectives* (Oxford: OUP, 2000), p. 40. Significantly, Hill does not say 'impossible'.

officers, the maxim 'If I am captured by enemies who will, by a brain scan, discover the combination of my missile launcher and destroy London, I will take a fatal pill'? The case is especially interesting because it shows that even perfect obligations can conflict—say, the obligations not to suicide and not to break promises to protect the population. It is important to see here that Kant's top-down method does not preclude approaching moral conflicts by seeking a principle that resolves them.[16] Our example identifies a *principled* exception—one warranted by a universalizable reconciling principle. *One may indeed consider such reconciling principles categorical imperatives with a small c.* Granted, Kant would countenance first-hand knowledge of duty in conflict cases—as opposed to mere good habit's yielding the right moral judgment—*only* in the light of a sense of some universal principle that covers the particular case. Reconciling principles may, then, be discoverable—or formulable—only by mature agents and perhaps only in rare cases of moral conflict. But the difficulty of discovery and formulation and the narrow scope of some reconciling principles are contingent matters; they are consistent with Kant's rationalism about the status of moral principles; and countenancing universalizable reconciling principles enables him to refute the charge of rigorism that has so long beset his ethical theory.

A different question is whether his theory allows countenancing incommensurability at all. Here we need two distinctions. One is between incommensurability and incomparability.[17] Affirming the former regarding (prima facie) duties denies a common measure of strength among them; affirming the latter regarding them denies the possibility of rational decision in some cases of obligational conflict. The second distinction is between incomparability and indeterminacy. The latter is the absence of any truth of the matter in certain conflict cases (including the truth that the conflicting duties are equally strong, since that would allow disjunctive final duties.) Both Kant and Ross reject incomparability in cases of moral conflict, but whereas Ross allows particularistic intuitive judgment as a rational basis of final moral judgment, Kant would demand *a generalistic basis*. The guiding idea is that the Categorical Imperative is the master principle and constitutes a basis for resolving any apparent moral conflict. Neither view, however, presupposes ontic indeterminacy in ethics. Granted, there is an innocuous indeterminacy whenever (as in certain dilemmas) two incompatible acts are both morally acceptable; but this may be dealt with on any plausible view either in terms of a need for some non-moral way to choose or by positing an overriding disjunctive obligation, e.g. to *A or B*.

The methodological differences between Kant and Ross should not be allowed to obscure a major concurrence—both respect moral common sense.[18] On either view, a

[16] Sometimes Kant even seems to presuppose that we always act on a maxim, as in fact the wording of at least the universality formulations of the Categorical Imperative may suggest.

[17] Much defense of a distinction between incommensurability and incomparability is provided by Ruth Chang, e.g., in her introduction to her edited collection, *Incommensurability, Incomparability, and Practical Reason* (Cambridge: Harvard UP, 1997). For critical discussion of the distinction and of whether incommensurability forces a recognition of a kind of relativism, see Martjin Boot, "Parity, Incomparability, and Justified Choice," *Philosophical Studies* 146 (2009), pp. 75–92 and *Incommensurability and its Implications for Practical Reasoning, Ethics and Justice* (London and New York: Rowman and Littlefield International, 2017).

[18] See on this point Karl Ameriks, "A Common-Sense Kant?", *Proceedings and Addresses of the APA* 79 (2005), pp. 19–45.

correct resolution is implicitly general: its basis is features of the conflict situation such that *any* situation having those features has the same resolution. Ross would think those features may holistically yield an intuitive judgment and need not provide an actual basis for universalization (though Ross might have held, as I do, that they commonly do provide it). Kant's view, however, requires that even an intuitive judgment must be so connected with the agent's understanding that an applicable adjudicatory maxim is both formulable and universalizable. Kant would require that, at least for someone of high-level moral understanding, there be a route from the bottom to the top.

How great a difference is this? Ross stresses the need for moral maturity as required for understanding moral principles and, implicitly, moral obligation; Kant might agree on its importance and indicate that moral imagination might also be required for formulation of an appropriate adjudicatory maxim. The recalcitrant difference is that Ross would allow intuition to justify a moral judgment even where formulating the relevant maxim is at best possible in principle, whereas Kant would require that the maxim must be in some way accessible to the agent and, perhaps, in some way serve as a premise in practical reasoning, even if not actually formulated. As a predominantly top-down theorist, Kant demands that moral agents be able to approximate a top-down justification of their (significant) actions by (potentially) framing a universalizable maxim for them, whereas an intuitionist will allow certain particularistic intuitions to carry sufficient epistemic autonomy for moral justification. This difference is significant enough to constitute one in ethical method, but not sufficiently deep to undercut the similarities that make both Kantian and intuitionist views rationalistic and anti-consequentialist.

4 Is Kant's method in ethics unqualifiedly top-down?

Ross did not note all the similarities between intuitionism and Kant's view pointed out here, but he did consider Kant's view "a version of intuitionism."[19] To what extent might this remark indicate not only elements of intuitionism in Kant's moral theory but also some easily overlooked commonality with one kind of bottom-up method?

That Kant relied on intuition, broadly conceived simply as a capacity for rational, non-inferential apprehension of a priori truths, need not be argued (at least not here). And in many places he makes judgments that apparently reflect moral intuitions even if he might defend them by derivation from the Categorical Imperative. The *Lectures* represent a good source of these. There, immediately after clarifying the abstract point that "moral goodness endues man with an inner, absolute worth," he says, "For example, the man who keeps his word has always an immediate inner worth of the free will, apart altogether from the end in view."[20] Another passage affirms the intuitive connec-

[19] See Sir David Ross, *Kant's Ethical Theory* (Oxford: OUP, 1954), p. 21. Ross should not be taken to ignore Kant's special use of 'Anschauung', commonly translated by 'intuition', but Ross might have noticed Kant's saying (in the context of discussing the Humanity Formula and its relation to dignity) that in indicating his reason for formulating the Categorical Imperative in different ways he intends "to bring an idea of reason closer to intuition (by a certain analogy) and thereby to feeling" (4: p. 436).

[20] Infeld trans. (cited in note 5), p. 5. Heath's translation has 'gives man an inner absolute worth of morality'. See Heath and Schneewind, op. cit., p. 44. In either case the striking point is that Kant speaks of the inner

tion between serving others and their incurring a debt of gratitude (p. 54). Apparent moral intuitions seem also to govern much of what Kant says about responsibility for consequences, for instance that "if I owe money and I do not pay my debt when I ought, then if my creditor goes bankrupt in consequence, it is my fault" (p. 59). A broader affirmation of what appears to be the power of rational intuition is "[C]ommon human reason … knows very well how to distinguish in every case that comes up what is good and what is evil …there is no need of science and philosophy to know what one has to do in order to be honest and good" (4: p. 404).

Other passages exhibit intuitive affirmations of intrinsic value, as opposed to obligation or responsibility, e.g. that "intellect has intrinsic worth irrespective of the purposes to which it is applied" (p. 135). The point here is not that Kant cannot argue for these points or even find principles that subsume them. The point is that here and in other passages (such as the section on duties to oneself, pp. 117–126) he makes apparently intuitional claims which represent apprehensions that are neither empirical nor contingent nor drawn from prior premises. Indeed, he seems to regard as universal what appears to be a kind of rational intuition: in the context of discussing conscience he says, "In regard to his natural obligations, nobody can be in error; for the natural moral laws cannot be unknown to anyone, in that they lie in reason for all…" (Heath and Schneewind, p. 133).

Normative intuitions also seem apparent when Kant indicates that certain considerations have a basic (hence non-derivative) relevance to making moral judgments: "Certainly our weal and woe are very important in the estimation of our practical reasoning… Man is a being of needs, so far as he belongs to the world of sense… (*Critique of Practical Reason*, 5: p. 62). He also acknowledges the value of love as guiding us: "It is a very beautiful thing to do good to men because of love and a sympathetic good will" (5: p. 82). To be sure, "this is not the genuine moral maxim of our conduct" (5: p. 82). My point is that Kant seems to take as intuitively obvious the normative point that love is a good thing.[21] Granted, all of his points just cited are about *kinds*; but they concern determinate kinds and are presented as knowable in the non-inferential way characteristic of intuitional cognitions with singular or, in some cases, general moral propositions as objects.

There is a difference between intuitions whose contents are provable and those whose contents are not. If self-evident propositions were all unprovable, one might take Kant's apparent appeals to intuition (in the contemporary sense of the term) that seem applicable to the self-evident to be viewed by him as unprovable. But I do not find in the relevant texts any such restriction of intuition to the self-evident, or of the self-evident or the a priori to the unprovable (nor is the unprovability view regarding self-evidence correct[22]). Nonetheless, we might still say that, for Kant, moral philosophy,

Footnote 20 continued
worth of persons not as brutely possessed by them but as given by—presumably in some way *grounded in*—a property (a kind of agential one). This is both plausible and consonant with some leading current views in the ontology of ethics.

[21] One might indeed wonder whether Kant took the Humanity Formula as a kind of secular version of the second Biblical love commandment, as is perhaps suggested by Kant's own reference to it in, among other places (5: p. 83).

[22] The error here may trace to assimilating the self-evident to the Aristotelian indemonstrable, roughly that to which nothing is epistemically prior. I have discussed the difference in chapter 2 of *The Good in the Right*,

as distinct from ordinary moral education, should begin with the *concept* of a rational being—which is very general and wider than that of a human being—and should yield categorical, necessary, a priori principles. The Categorical Imperative is the master principle; but, once agents grasp it, they can solve moral problems by appeal to lower-level principles and even concrete cases exhibiting generalizable elements. Final moral authority resides higher up, ultimately at the top level; but the role of intuitions and of bottom-up thinking may be a starting point for determining an appropriate principle, and intuitions may have heuristic importance and even (derivative) evidential weight.[23]

One further dimension of methodological comparison needs clarification. It might seem that Kant's mainly top-down method takes sound moral thinking to be both subsumptivist regarding determination of moral judgment and intellectualist regarding moral thinking. For *subsumptivism*, singular moral judgments reflect the agent's applying a moral generalization to the action(s) judged, even if without consciousness of doing this. For *intellectualism,* sound moral judgment must arise from a reasoning process—a kind of exercise of intellect. Let me take these views in reverse order.

The famous four examples in the *Groundwork* easily suggest that the agent judges that (e.g.) making a lying promise is wrong by reasoning to that judgment from the Categorical Imperative. But surely the role of the Imperative in those passages is (or might be) to *test* a projected judgment or to empower us to *know* what is right; simply *arriving* at the right judgment given temptation does not require such reasoning. Granted, for Kant the reasoning must be appropriately *accessible* to agents, and we might take its content to represent the agent's cognitive stance at the time —a structural position disposing one to reason if the occasion calls for it—but that is a weaker requirement. The shopkeeper example seems to illustrate the sufficiency of the accessibility requirement for the possibility of acting on a universalizable appropriate maxim (4: p. 397). As with linguistic rules, we can *abide by* a categorical imperative without *reasoning from it*, much as we can act under an authority without appealing to it.

If, for the reasons given, Kant's method does not commit him to intellectualism, then we should not take it to imply subsumptivism. Subsumptivism is not necessary or sufficient for intellectualism, but it is a main case of it that may seem unavoidable for a top-down theorist. Moreover, we can distinguish two kinds: *justificatory subsumptivism* is the epistemological view that sound singular moral judgments must be *capable of* justification as instances of a moral "law"; *genetic subsumptivism* is the psychological view that such judgments must be *arrived at* by way of a generalization.

Footnote 22 continued
but here I would stress simply that Kant's ethics does not depend on any such conflation. He does say that practical laws (in the form of universalizable maxims) are "*indemonstrable* and yet *apodictic*" (*Doctrine of Virtue* 224], but here he may well be suggesting both that our freedom sometimes allows more than one law as providing a directive for action in a given context or that its rational support from the Categorical Imperative is non-demonstrative—even if synthetic a priori—thus in either case including what would now be called intuitive. The element here of (limited) pluralism about rules of action—often thought impossible for Kant—deserves serious study but cannot be pursued here.

[23] It may be useful to recall Aristotle's distinction between "arguments from principles and arguments toward principles" (NE 1095a31ff.) Even if the top level has higher epistemic authority, one might countenance, as Aristotle apparently does, some degree of epistemic authority on the part of the cases from which (as with intuitively clear concrete instances) one might, perhaps by intuitive induction, argue for a principle or explanation.

Kant is apparently committed to the first but not the second, intellectualist version. He indeed makes room for habits and other elements in character to yield sound moral judgment *spontaneously*. If he is intellectualist, it is about moral knowledge, not about (elementary) moral education or the possibility of morally creditworthy action.

5 Ethical method and normative ethics

The theoretical importance of methodological considerations for moral philosophy is perhaps beyond doubt, and this holds for the distinction between top-down and bottom-up methods. But we might also fruitfully consider the normative ethical significance of Kant's method in ethics.

Significantly, Kant seems to take as axiomatic that good will is "good without qualification" (*Groundwork*, 4: p. 393). He does not rule out other goods' being inherently (hence non-instrumentally) good. Indeed he includes, as good, talents of the mind and qualities of temperament widely considered virtuous. If none of these is good in itself, it is at best difficult to see why he should say—what may be taken to express an intuition of *unfittingness*—that "a rational and impartial spectator can never feel approval in contemplating the uninterrupted prosperity of a being graced by no touch of pure and good will…" (4: p. 393). Arguably, Kant could be applying this point only to the gifts of fortune, which might be considered only instrumentally good; but even in that sentence he speaks of happiness in a way that encompasses goods that are not just instrumental. In any case, supposing Kant did not take good will to be to the only thing good in itself, he might have considered it the *highest* good—an assumption which allows for other non-instrumental goods.

My positive interpretation is that good will has at least three essential characteristics: it is good in itself; it cannot be undeservedly possessed;[24] and it is "the indispensable condition of our worthiness to be happy" (4: p. 393). In *both* the second and third points, good will differs from every other inherent good. All three points seem plausible. They are difficult to show, but my concern is to indicate that Kant apparently does not argue for them beyond giving illustrative examples. He perhaps finds it intuitive that there is no other *conceivable* unqualifiedly good entity. If we have here a postulate—as opposed to a generalization—at the top, it is surely one that manifests a respect for a priori intuition.

What this passage and much of the *Groundwork* suggest is that Kant's method is not only *generalizationally top-down*—giving a central place to a master principle and principles deducible from it—but also *valuationally top-down*: giving a central place to good will as having absolute value and *conditioning*—though not *instrumentalizing*—the value of all else in our lives.[25] Other intrinsic goods are conditioned in the

[24] I have discussed the valuational status Kant attributes to good will in a number of places, arguing that, by contrast with other inherent goods, it cannot be undeservedly possessed. See ch 3 of *Practical Reasoning* (London: Routledge, 1989) and, for more detail, in a successor volume, *Practical Reasoning and Ethical Decision* (London: Routledge, 2006). Granted, one could say a *deserved* honor cannot be undeservedly possessed; but that trivial point is consistent with the status of good will described in the text in relation to other intrinsic goods.

[25] For detailed discussion of the sense in which good will conditions other goods, see Eric Watkins, "The Unconditioned Goodness of the Good Will," forthcoming.

sense that without good will one is not worthy to possess them; they are not thereby rendered mere means to good will—they need not even conduce to it. When we ask what good will is, however, we are told too little, at least in the Preface: it is "good through its willing alone" (4: p. 394). I find this plausible. In contemporary language, good will is a construct out of good intentions: what, referring to good will, he calls "its willings." How can we determine the crucial kinds of intentions? By appeal to the Categorical Imperative.

Kant saw a problem here. In one place he notes the *paradox of method*: "…that the concept of good and evil is not defined prior to the moral law, to which, it would seem, the former would have to serve as a foundation; rather, the concept of good and evil must be defined after and by means of the law" (5: pp. 62–63; cf. 4: p. 450, which acknowledges "a kind of circle"). As he says in an earlier passage, "The *autonomy* of the will is the sole principle of all moral laws, and of the duties conforming to them" (4: p. 33). Kant presumably took the paradox to be resolved by the way in which a proper use of the Categorical Imperative leads to determining what motives (and types of conduct) are morally good (or morally bad). But its success in doing this appears to depend on resolving problems concerning how to formulate maxims and determine which of them may be universalized.[26] Discussing these problems would take us too far from my main task, but their magnitude is suggested by clear difficulties in Kant's treatment of his four examples in the *Groundwork*. Here I will simply say that there is no reason Kant cannot be understood as wishing us to be guided in part by the Humanity Formula in ascertaining whether a projected deed conforms to universalizability formulations of the Categorical Imperative.[27] How might such guidance work?

Recall the Humanity Formula: "So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" (4: p. 429).[28] In relation to this Kant says, "The ground of this principle is: *Rational nature exists as an end in itself,*" and from the principle "it must be possible to derive all laws of the will" (4: p. 429). The *Groundwork* gives a quite abstract and mostly negative characterization of what it is to treat persons as ends, and the literature on Kant's ethics contains much intended to explicate what he means and to show how the Categorical Imperative can do the normative work he intended for it.[29] I doubt that it can do this work unless given an interpretation that, though consistent with his

[26] J. S. Mill raised one kind of problem of this kind in ch 1 of *Utilitarianism,* and there have been a host of objections raising such problems since Mill.

[27] This is not to suggest that Kant provided enough explanation of the Humanity Formula to solve the problem. However the problem is described, I agree with Karl Ameriks that "Ultimately…there remains in Kant a central and insufficiently justified belief in an intrinsic connection between morality and absolute freedom." See *Interpreting Kant's Critiques* (Oxford: OUP, 2003), pp. 161–192, 162.

[28] Here I use Mary Gregor's translation (romanizing the italics); she uses 'merely' where some others have 'simply', which the context indicates is less close to Kant's overall intentions.

[29] Here Hill speaks for the constructivist line of interpretation of Kantian ethics favored by John Rawls and many of his students: "To review, on the Kantian perspective, the ultimate source of human values is not Platonic forms, natural teleology, God's will, or universal human sentiment. Ultimately, all that is valuable for us stems somehow from the reflective endorsements of human beings" (op. cit., p. 77.) Hill's rejected alternatives do not include rational intuition. By contrast, Allen W. Wood holds (consistently with my view) that "Kant agrees with the British rationalists' endorsement of the idea that values lie in the nature of things rather than being conferred on things by someone's will." See *Kantian Ethics* (Cambridge UP

overall ethics, is not part of his explication of either the Humanity Formula or any other formulation of the Categorical Imperative. If, however, one understands treatment of persons in the light of the overall tenor of his ethical writings and in relation to his examples, his ethical theory may be regarded as both plausible and powerful.[30]

The significance of the *Kantian circle problem*, as we might call it, can be seen by noting a certain irony in Kant's reaction to Hume. In his ethics, Kant has apparently not, in his concern to transcend Hume (among others), fully applied the power of the synthetic a priori, a power he enlists against Hume in his theoretical philosophy. Hume is (in parts of his work) a skeptic about normativity as he is about causation and much empirical knowledge. For Hume—insofar as he is a thoroughgoing instrumentalist—since moral judgments motivate us and judgments based on reason do not, the former are not so based. Reason, then, yields no categorical imperatives or valuational grounds for them, as it does for Kant; it is *content-neutral*: it never tells us what is worth desiring, but at most how to satisfy our "basic" (non-instrumental) desires.

If one reads Kant as providing no substantive values or determinate obligations—e.g. no definite content constraining the good will—he comes perilously close to Hume in apparently seeking to provide an ethic with content-neutral foundations. If, e.g., treating others as ends is *just* treating them in ways they can rationally agree to *apart* from any specific values or rational desires—where valuation and rationality are content-neutral—then too much is left open. Similarly, restriction to what one can "reflectively endorse"—unless constrained by more than logic and analytic truths—also leaves too much open.[31] If no substantive normative standards guide the endorsement, coherence (a structural property entailing no definite content) is apparently all reason can provide as a constraint.

Suppose, however, Kant presupposes—as so many of his points about beneficence, need, and happiness suggest—some positive and negative intrinsic values other than, but in part guiding, good will. This may be implicit in his view that "The concepts of good and evil first determine an object for the will" (5: p. 67). Making substantive axiological presuppositions does not prevent his developing his ethics as he apparently intends—giving autonomy a central role while avoiding circularity. He appears in any case to consider certain affirmations of intrinsic value to be synthetic a priori, and he has no less conviction regarding the wrongness of, for instance, lying and ingratitude. Even in treating good will as the only unconditional good, he acknowledges that it

Footnote 29 continued
2008), p. 142. Some such axiological realism certainly seems explicit in 4: p. 428, where Kant ringingly affirms the existence of something ("the human being") "*which in itself* has an absolute worth."

[30] Much that I say here and in the remainder of this section is controversial, but I provide a full-scale interpretation of the Humanity Formula—taken in an intuitive way—in *Means, Ends, and Persons*: *The Meaning and Psychological Dimensions of Kant's Humanity Formula* (OUP 2016), and I there suggest how endorsing at least much of its normative force is consistent with Kant's position. My effort here is less ambitious, emphasizing only a few elements of a Kantian axiology. If we emphasize our own perfection and the happiness of others as what Kant considers obligatory ends, then the content of those notions provides teleological grounds for moral obligations. For development of this line see Wood, op. cit., esp. chapter 9, on duties, and also Huston Smit and Mark Timmons, "Kant's Grounding Project in *The Doctrine of Virtue,"* in Mark Timmons and Sorin Baiasu, eds., *Kant on Practical Justification* (Oxford: OUP, 2013), pp. 229–268.

[31] This is a major claim I cannot defend here. But I have suggested a related problem for constructivism as Rawls and others have presented it: it appears, ironically, to be best conceived as committed to reductive naturalism. See my "Moral Knowledge," in Sven Bernecker and Duncan Pritchard, eds., *The Routledge Companion to Epistemology* (London: Routledge, 2011), pp. 380–392.

"need not be the sole and complete good" and that there are others (4: p. 396). He might consistently grant that some of these other goods, such as "intelligence, wit, judgement … or courage, resolution, and constancy of purpose" (p. 393) can normatively ground reasons for action and can be seen a priori to do so.

Must Kant see the determinate axiological and deontic commitments needed to give basic direction to good will as introducing impurities in practical reason? Even a strongly top-down theorist need not be this kind of purist, and Kant need not take synthetic a priori propositions, or even self-evident propositions, to occur only at the top-level of his valuational hierarchy. He can consider good will the only *unconditioned* good even if he considers it broadly constrained by taking account of lower-level *intrinsic* goods. This need not put the bottom above the top—the level of intrinsic goods is one of a priori discernment of necessary truths with normative implications for action—nor need we take the good to be more basic than the right, or subordinate the will to consequentialist standards.

Here I suggest a possibility for future reflection, one that accords with a top-down methodology that places bearers of intrinsic value at the top and construes as a priori the idea that what has dignity has intrinsic value. Suppose that dignity—a high axiological status of persons—is at least largely grounded in autonomy, as Kant apparently holds.[32] Then certain substantive moral principles are derivable. One is that it is wrong to impair agency, say by coercion. This in turn both implies and partly explains why it is wrong (in certain cases) to cause pain in persons. This deontic "theorem" is defensible on two apparently a priori assumptions: that, as requiring attention, or at least space, in the limited consciousness of (psychologically) finite beings, pain tends to impair their autonomy, and that impairment of agency, in turn, reduces autonomy, at least in scope if not in actual exercise. Reducing autonomy is inherently bad and can ground of prohibitional principles. A further, subtler reason for the (prima facie) wrongness of causing pain is that, in Kantian terms, by virtue of the great aversive motivational power of pain in finite rational agents like human beings, pain tends to induce heteronymous motives. People tend to go to drastic lengths to escape it.[33]

A related point drawing specifically on the central concepts in the Humanity Formula (which we might in any case consider important in understanding the other formulations of the Categorical Imperative) is this: it is an a priori truth, and certainly intuitive, that one's treating anything as an end requires *caring* about it non-instrumentally and, if it is a person, further requires conceiving something or other as *its good*. We cannot act for the *sake* of persons—a possibility entailed by their being ends in themselves—unless we take them to have a good, in a sense entailing that certain things are *in their interest.* Similarly, avoiding treating something merely as a means a priori entails not relating to it merely instrumentally—as a thing that can be

[32] Kant says, e.g., "the dignity of man consists precisely in his capacity to make universal law" (440, p 107), where "the mere dignity… of rational nature in man… should function as an inflexible precept of the will" (85, p. 106).

[33] This defense of Kant against the empty formalism objection usefully contrasts with Christine Korsgaard's idea that "our own perfection and the happiness of others are identified as obligatory by his contradiction in the will test." See "Acting for a Reason," in her *The Constitution of Agency*" (Oxford 2008), pp. 205–229, 220). My view is compatible with one way to interpret that test but avoids dependence on contingent or highly controversial premises. For a detailed discussion of how moral obligations are determined in Kant's ethical framework, see Smit and Timmons, op. cit.

discarded after fulfilling its instrumental purpose. The notions of ends and means in question impose substantive a priori requirements on the content of good will. One important element is what is best understood as equal moral status among rational beings: all of us are ends with the dignity of autonomy; hence none may be treated merely instrumentally. The former conception requires a broadly egalitarian kind of justice: for all of us, as ends, it both prohibits causing pains that impair autonomy and implicitly entails seeking the (clearly diverse) goods that preserve and enhance autonomy. The latter imposes duties to avoid "using" people merely as means—which encompasses multifarious universalizable prohibitions.

If the reflections of this paper are sound, we may view Kant as a methodologically top-down thinker in his metaethics, a rationalist in his moral epistemology, and a master principle theorist in his normative ethics. But we need not take him to be committed to either a strongly intellectualist or a subsumptivist moral psychology. We cannot consider him an intuitionist, but if ethical intuitionism is in part constituted by attributing epistemic authority to certain intuitions regarding certain kinds of particular cases, Kant may be qualifiedly considered intuitionistic, as perhaps Ross perceived. This holds of Kant insofar as the intuitions in question concern *types* of actions, which are abstract—e.g. lies, broken promises, and certain restrictions of liberty. This apprehensional element in intuitionism, however, leaves open whether the *degree* of epistemic authority of such low-level intuitions is as great as that of intuitions or other cognitions occurring higher up. On that score, Kant's generalism separates him from Ross and other intuitionists. Nothing said here, however, precludes proponents of Kantian ethics from at least largely accepting some important elements in what I call *Kantian Intuitionism*[34]—though it might also be called an intuitionist Kantianism—which integrates a contemporary version of a Rossian intuitionism with an interpretation of the Humanity Formula that, though containing much that Kant did not explicitly hold, is consistent with his major points about the Categorical Imperative as expressed in that formula and with value commitments he expresses in many passages and with his emphasis of dignity as essential to rational persons. Perhaps the best way to understand Kantian ethics as a whole is to conceive the summit from which he abstractly views the normative territory as integrated with some of the intuitive insights found mainly, and perhaps only, at the middle and bottom levels where much is in clear intuitive view.[35]

[34] I have articulated this position in ch 3 of *The Good in the Right: A Theory of Intuition and Intrinsic Value* (Princeton, PUP, 2004) and proposed extensions and refinements of it in "Kantian Intuitionism as a Framework for the Justification of Moral Judgments," *Oxford Studies in Normative Ethics* 2 (2012), pp. 128–151. My understanding of the Humanity formula is developed at length in *Means, Ends, and Persons: The Meaning and Psychological Dimensions of Kant's Humanity Formula* (Oxford: Oxford University Press, 2016).

[35] This paper has benefited from presentations at Birkbeck College London, Colgate University, Colorado State University, the Goethe University, Frankfurt, and St. Louis University, and for helpful comments I particularly want to thank Gabrielle Gava, Patrick Kain, Pauline Kleingeld, Mark Timmons, Eric Watkins, Marcus Willaschek, and, especially, Karl Ameriks and Katharina Kraus.