The Morals of Metaphysics: Kant’s *Groundwork* as Intellectual *Paideia*

Ian Hunter

In fact this satisfies a longing for the transcendent, because in so far as people believe they can see the ‘limits of human understanding’, they believe of course that they can see beyond these.

Work in philosophy . . . is really more work on oneself.
—Ludwig Wittgenstein, *Culture and Value*

To approach philosophy as a way of working on the self means to begin not with the experience it clarifies and the subject it discovers, but with the acts of self-transformation it requires and the subjectivity it seeks to fashion. Commenting on the variety of spiritual exercises to be found in the ancient schools, Pierre Hadot remarks that:

Some, like Plutarch’s *ethismoi*, designed to curb curiosity, anger or gossip, were only practices intended to ensure good moral habits. Others, particularly the meditations of the Platonic tradition, demanded a high degree of mental concentration. Some, like the contemplation of nature as practiced in all philosophical schools, turned the soul toward the cosmos, while still others—rare and exceptional—led to a transfiguration of the personality, as in the experiences of Plotinus. We also saw that the emotional tone and notional content of these exercises varied widely from one philosophical school to another: from the mobilization of energy and consent to destiny of the Stoics, to the relaxation and detachment of the Epicureans, to the mental concentration and renunciation of the sensible world among the Platonists.

While successfully applied to ancient philosophy, this approach has not been widely exploited in the history of philosophy more broadly. There is,
However, at least one study of medieval metaphysics in these terms,\(^3\) and there are some important discussions of early modern Stoicism and Epicureanism.\(^4\) And a recent study of Hume shows the fruitfulness of the approach for Enlightenment philosophy.\(^5\) It is all the more surprising then that there seems to have been no serious attempt to approach Kant’s moral philosophy in this way.

Hadot and Foucault seem to have felt that the abstract and academic character of modern philosophy meant that it was no longer cultivated as a way of life—this despite Foucault’s treatment of Descartes’s *Meditations* as a spiritual exercise designed to allow the mind to achieve certainty by inducing, then overcoming, skepticism.\(^6\) The greater obstacle to approaching Kant’s moral philosophy as a way of life, however, comes from the fact that both its friends and its enemies insist on its formal (or formalistic) character. American Kantians thus take it for granted that Kant’s moral philosophy represents, not the cultivation of a moral life, but the formal recovery of the rational grounds that make life moral.\(^7\) As far as its Thomistic and communitarian opponents are concerned, this formalism is the ruin of Kantian ethics, uprooting its judgments from moral tradition and detaching them from the moral community whose substantive virtues provide the ground and purpose of morality.\(^8\) In either case, whether we view it as

---


**Ian Hunter** is Australian Research Council Professorial Fellow in the Centre for History of European Discourses, University of Queensland. He is the author of several works on early modern political and moral thought, including most recently *Rival Enlightenments: Civil and Metaphysical Philosophy in Early Modern Germany* (2001). With David Saunders he is coeditor of *Natural Law and Civil Sovereignty* and of a new edition of Andrew Tooke’s first English translation of Samuel Pufendorf’s *De officio, the Whole Duty of Man*, both of which are to be published in 2002.
the rational foundation or as the rationalist deracination of moral life, we are prevented from approaching Kant’s formal philosophy as itself a moral culture of a particular kind.

In order to break out of this weary standoff it is necessary to radically reorient our approach to Kant’s moral philosophy. We must learn to see its formal purity, not in terms of the pursuit of rational grounds, but as an aspiration arising from the incitement to and cultivation of a certain kind of moral purity. This viewpoint cannot be reached by asking the familiar questions: What is Kant’s pure moral law and how is it known and validated? Does Kant rely solely on the rational purity of the moral law in making judgments, or does he also allow the feelings and inclinations to play a part? Can morality be founded in formal insight into rational grounds or does it require the cultivation of moral character and the acknowledgment of moral community? Instead, if we are to acquire the level of detachment needed to understand the manner in which Kant’s philosophy takes hold of us, we must learn to ask a different kind of question: What is it that first leads us to turn to ourselves in expectation of finding within the commanding presence of a pure moral law? How do we first come to think of ourselves as beings divided between the freedom of a pure intellect and the desires of a sensuous nature? What must we do to ourselves—performing what inner exercises using what intellectual instruments—to acquire the deportment of someone who hears and obeys the commands of a higher rational self? And what is the source of the extraordinary spiritual prestige surrounding this deportment? In what follows I show why these questions are worth asking by providing indicative answers to them in a brief re-description of Kant’s *Groundwork of the Metaphysics of Morals.*

**The Way In**

Despite the remarkable lack of commentary on it, the fundamental relation between formal purity and a specific culture of moral purity finds symptomatic expression at the beginning of the *Groundwork* in the preface. Paradoxically, this occurs in the very formulation where Kant seeks to free a pure moral philosophy—the metaphysics of morals—from all dependency on man’s empirical moral nature and its discipline, moral anthropology:

Since my aim here is directed properly to moral philosophy, I limit the question proposed only to this: is it not thought to be of the utmost necessity to work out for once a pure moral philosophy, completely

---

cleansed of everything that may be only empirical and that belongs to anthropology? For, that there must be such a philosophy is clear of itself from the common idea of duty and of moral laws. Everyone must grant that a law, if it is to hold morally, that is, as a ground of an obligation, must carry with it absolute necessity; that, for example, the command “thou shalt not lie” does not hold only for human beings [Menschen], as if other rational beings [verünftige Wesen] did not have to heed it, and similarly with all other genuine moral laws; that, therefore, the ground of obligation must not be sought in the nature of the human being or in the circumstances of the world in which he is placed, but solely a priori in concepts of pure reason.10

Quite unexpectedly—given the standard readings—Kant motivates the need for a formal and universal moral philosophy via the idea that the universe of rational beings outstrips the world of humans. This means (“therefore”) that the grounds of moral obligation must be sought, not in a merely human nature, but in the formal or a priori concepts of pure reason suited to (transhuman) “beings of reason.”

Despite Kant’s attempt to distinguish “pure moral philosophy” from all moral anthropology—that is, from the repository of human figurations (Stoic, Platonic, Epicurean, Christological) used to configure personhood—it seems clear that this distinction is itself dependent on one such anthropology. We can discern this anthropology in the figuration of humans as a particular species of rational being (Vernunftwesen). This species is characterized by the union of a rational (verünftige) nature—shared with God and the angels—with a sensible (sinnliches) nature, consisting of man’s sensory faculties and sensuous inclinations. It is just this Christian-Platonic figuration of man as a rational being mired in the spatiotemporal world by his senses, and in the prudential world by his sensuous inclinations, that allows Kant to separate the metaphysics of morals from “empirical” anthropology. By tacitly invoking this metaphysical anthropology Kant can identify metaphysics with pure (nonspatiotemporal) insight into a moral law binding on a universe of pure intelligences only some of whom are human. This allows him to relegate all other anthropology to the “pragmatic” task of refining man’s sensuous inclinations to render them capable of receiving the pure moral law in the impure empirical world.11


11. For discussions of Kant’s “pragmatic” anthropology—that is, for discussions that ignore the role of his metaphysical anthropology in shaping the moral law itself—see Wood, “Unsociable
Arrived at in this manner, the metaphysical pursuit of formal foundations for morality itself takes on a profoundly moral character. For it holds the key to man’s participation in the world of rational beings to which he is drawn by the higher (intellectual) part of his own double nature. Conversely, in this setting, empirical moral anthropologies—that is, all the anthropologies Kant construes as dealing with merely human nature—are not just philosophical mistakes. Rather, they are seen as morally corrupting, miring rational beings in their human or sensuous natures, and thereby hindering their metaphysical refinement:

A metaphysics of morals is therefore indispensably necessary, not merely in order to investigate, from speculative motives, the source of the practical principles that lie a priori in our reason, but also because morals themselves remain subject to all sorts of corruption as long as we lack the guiding thread and highest norm for their correct judgment. . . . Now the moral law in its purity and genuineness . . . is to be sought nowhere else than in a pure philosophy; hence, this (metaphysics) must come first, and without it there can be no moral philosophy at all. That which mixes up these pure principles with empirical ones does not even deserve the name of philosophy . . . much less does it deserve the name of a moral philosophy, since by this very confusion it actually damages the purity of morals themselves and acts against its own end. [AK, 4:389–90; PP, pp. 45–46]

Kant’s opening separation of the metaphysics of morals from moral anthropology is thus something far more consequential than the meta-ethical distinction between an objective and a subjective, or a pure and applied, ethics. In positing it as the only discipline capable of perfecting man’s highest or noblest part—the pure intelligence that he shares with other beings of reason—Kant is not introducing metaphysics as a defeasible theory of the moral subject. Rather he is presenting it to his students and readers as the only discipline capable of purifying their sensuous natures and permitting their participation in the world of pure, self-governing intelligences. In other words, Kant is presenting metaphysics itself as the discipline of a prestigious way of thought and life. If this is so, then the metaphysics of morals will turn out to be grounded in the morals of metaphysics.

If Kant opens the Groundwork by embedding the formal or pure philosophy of morality in the desire for (metaphysical) moral purity, that is
because his figuration of man as “sensibly affected rational being” was the latest incarnation of a special moral anthropology—that of university metaphysics—whose function is to incite this desire. Deeply rooted in the history of Christian Platonism, the figuration of man as a pure intellect mired in a sensuous nature was installed in the medieval university by Albert the Great, elaborated by his student Thomas Aquinas, and transmitted to the modern philosophy faculty via the line running from Leibniz through Wolff to Kant. Given the centrality of metaphysical anthropology to Kant’s entire undertaking, we may ask why its role as the enabling condition of a metaphysics of morals has attracted so little commentary. The answer, we may conjecture, is that most Kant commentary is written by those who have entered philosophy through the figure of Homo duplex transmitted in university metaphysics, that is, by those who have learned to feel that Kant’s metaphysics of morals provides the only path to moral purity, against the odds of their own sensuous natures. John Rawls, for example, assumes that Kant’s moral law applies to “the normal conditions of human life,” which he then specifies in terms of us being “reasonable and rational persons endowed with conscience and moral sensibility, and affected by, but not determined by, our natural desires and inclinations.” In formulating the tasks and limits of moral philosophy, Christine Korsgaard also takes the metaphysical anthropology for granted: “The moral law commands you to seek your own moral perfection: the holiness of your will. This cannot be achieved in the course of your life, for no one with a sensuous as well as a rational nature has a morally perfect disposition.” For his part, Allen Wood claims that Kant’s version of Homo duplex amounts to a “controversial empirical thesis about human nature,” but then he proceeds to use the tension between man’s pure rational nature and impure empirical inclinations as the framework for his own discussion of the thesis.

The figure of two-natured man, however, is neither a formal postulate nor an empirical hypothesis, but something else altogether: the cultural device through which those being initiated into a particular moral “school” are induced to take up a certain relation to themselves as the condition of commencing a work on the self. Wood comes close to realizing this in his comment that “Kant’s moral principles and his theory of human nature are

15. Wood, Kant’s Ethical Thought (Cambridge, 1999), p. 8; hereafter abbreviated KET.
designed only to add to our discontent with ourselves. . . . Kant thinks that as rational creatures our condition must be one of dissatisfaction, self-alienation, and endless striving” (KET, p. 334). Yet he squanders this insight by treating the self-discontent that Kant’s anthropology is designed to incite as if it were justified by the actual moral nature—torn between pure moral reason and wayward sensuous inclinations—that Wood is convinced we have. As Hadot’s comments suggest, however, we may be discontented with ourselves in several different ways, depending on what dimension of our lives a particular moral anthropology raises to the threshold of moral concern and configures for moral transformation. We grasp the plurality of paths to moral subjection, and the rivalry between them, by recalling that the early modern “civil” philosophers—Hobbes, Pufendorf, Thomasius—self-consciously rejected the metaphysical anthropology of homo duplex. Regarding the deportment it formed as hostage to both inner illuminism and clerical supremacism, they adopted in its place a quasi-Epicurean anthropology of man as a dangerous creature of his passions, treating this as the only one suited to an ethics of civil decorum backed by political constraint.16

Without realizing it, by identifying the metaphysical anthropology of homo duplex with the moral subject as such, modern Kantians betray their prior induction into a specific practice of self-cultivation and their partisanship for a historically contested moral culture. Once this identification has been accepted, it becomes impossible to investigate Kant’s moral philosophy as a particular kind of moral culture or way of life, for now one sees oneself as a “sensibly affected rational being” and begins to conduct one’s life in the manner of a Kantian, that is, to aspire to a certain kind of intellectual purity through the purifying effects of Kantian philosophy itself. Through our preliminary account of the dependency of “pure moral philosophy” on the culture of moral purification transmitted by university metaphysics, we have opened up a different way of understanding Kant’s moral philosophy in the _Groundwork_. We shall approach the _Groundwork_, not as a theory of the moral subject, but as a repository of devices for inducting students into the cultivation of a prestigious moral self.

**Section 1: Inducing the Desire for Philosophy**

In purporting to show that the “supreme moral principle” is already present in “popular moral consciousness”—and hence may be recovered through philosophical analysis—section 1 of the _Groundwork_ takes the

---

reader through a series of related arguments. First, Kant claims that his readers already know that the only unconditionally good thing is a good will. They know that the good will is an incomparably higher good than all the ends we associate with happiness—“Power, riches, honor, even health”—and all the virtues to which the pagan philosophers aspired: “Moderation in affects and passions, self-control, and tranquil reflection” (AK, 4:393, 394; PP, pp. 49, 50). Next, Kant moves to elucidate this still somewhat esoteric conception of the good will by showing that it is already contained in the popular idea of doing one’s duty for its own sake (see AK, 4:397–400; PP, pp. 52–55). He then argues that his concept of duty must be understood as the determination of the will through the mere idea or thought (Vorstellung) of duty (see AK, 4:401–2; PP, pp. 56–57). Finally, Kant concludes that in constructing this conception of the moral principle he has done nothing more than clarify a principle already present in ordinary moral consciousness (see AK, 4:403–5; PP, pp. 58–60).

To the extent that it is provided by scholars who have entered moral philosophy through the Kantian anthropology, modern commentary generally takes Kant at his word, treating the arguments of section 1 as an analytical attempt to uncover the conceptions of the good and of moral obligation already contained in “ordinary moral consciousness.” On the face of it this is an extraordinary way of proceeding, in part because it seems to hinge on getting the bearers of this consciousness (Kant’s students and readers) to declare its contents and in part because what they are supposed to affirm is itself so extraordinary: not, as one might expect, such goods as health, wealth, or power, and not such virtues as fortitude, compassion, or inner tranquility, but the goodness of a will whose purity consists in being aimed at none of these things. This way of proceeding becomes far less extraordinary, however, as soon as we recall the metaphysical anthropology that frames it. After all, it is just through this anthropology that Kant’s students and readers come to think of themselves as beings whose ordinary self—the one that pursues the merely material goods of health, wealth, compassion, tranquillity—obscures a latent higher self whose goodness consists solely in the purity of its willing. Students and readers who have been initiated into this way of relating to themselves will not only declare that their consciousness harbors such a higher concept of morality, but they also will treat this declaration as part of the culture of self-purification through

17. See C, pp. 55–67; Karl Ameriks, “Kant on the Good Will,” in Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar, ed. Otfried Höffe (Frankfurt am Main, 1989), pp. 45–65; and KET, pp. 17–49. Although, note Wood’s acknowledgment that “some of these claims clearly go well beyond anything Kant can pretend to draw solely from common rational moral cognition. They involve Kant’s theory of human nature” (KET, p. 25).
which they seek to groom themselves in the image of this concept. Seen in this light, the arguments of section 1 reveal a character quite unlike the analytical recovery of concepts already given in experience.

Kant’s opening appeal to the reader’s “existing” knowledge that the only unconditional good is a good will gains its force, not from conceptual analysis but from a quite different source: an evocative pedagogical presentation of the superiority of the contemplative over the prudential way of life. In declaring that the ends of civil happiness and the virtues of self-control are not the highest good, Kant grounds his affirmations, not in any argument against these rival ethical doctrines—doctrines in fact espoused by the Early Modern civil philosophers—but in the spiritual prestige already attaching to the ideal of contemplative autarky or intellectual autonomy.\(^{18}\) Kant’s “anticongquentialism” thus flows directly from the metaphysical anthropology, which posits rational being’s capacity to will independently of all external sensible goods or ends, purely to realize its own rational nature, thereby making the rational will “good in itself”: “A good will is not good because of what it effects or accomplishes, because of its aptness for attaining some proposed end, but simply through its willing; that is, it is good in itself and, beheld for itself, is of incomparably greater worth than anything it could bring about merely in favor of some inclination or, if you like, the sum of all inclinations” (*AK*, 4:394; *PP*, p. 50).

In this setting, the inferiority of empirical prudential ethics appears to arise from its pursuit of happiness in the “external” goods found in the world of space, time, and utility—goods lacking unity because of the variety of man’s sensuous inclinations and lacking certainty because of their need for actualization outside the self. The superiority of the good will, however, arises from the fact that, “beheld for itself,” it is freed from all empirical outcomes, thereby obtaining the autarky that in fact constitutes goodness for the contemplative ethos: “Even if . . . this will should wholly lack the power to carry out its intentions—if with its greatest efforts it should yet achieve nothing and only the good will were left—then it would still shine like a jewel for its own sake, as something containing its entire worth in itself” (*AK*, 4:394; *PP*, p. 50). In evoking the figure of the autarkic rational will, therefore, Kant is not reminding his readers of something of which they are already obscurely aware. Rather, he is offering them an image of the exalted personage they might become if only they will turn away from external prudential concerns with “power, riches, honor, even health” and begin the speculative purification of their inner wills. He is offering them a spiritual inducement to relate to their moral self in a new way, to reorganize

---

their moral life around the inklings of a moral will of which they were scarcely aware but on whose purity their moral future now hangs.

In this light, Kant’s appeal to his students’ sense of “duty for its own sake”—formed no doubt in religious, military, and pedagogical institutions requiring unconditional obedience—is no simple elicitation of evidence for the moral law’s preexistence. In fact it is a means by which his students can be induced to subject themselves to the law as something that already commands them from within. The crucial thing to note in this regard is Kant’s initial characterization of duty: “We shall therefore take up the concept of duty, which contains that of a good will though under certain subjective limitations and hindrances, which, however, far from concealing it and making it unrecognizable, rather bring it out by contrast and make it shine more brightly” (AK, 4:397; PP, p. 52). Here Kant provides his students with a new way of relating to their (still unfocused) sense of duty. By treating its compulsive character as arising from the form in which a pure rational will encounters the “subjective limitations and hindrances” of their sensuous natures, Kant incites his students to view their ordinary sense of duty (no matter what its source) as if it were their dimly “sensed dependency” as material beings on the self-governing community of intelligences in which they participate as immaterial (rational) beings.

Kant’s initial formulation of the unconditional or categorical character of the moral law—“so [setting aside inclinations] there is nothing left to determine the will except objectively the law and subjectively pure respect for this practical law” (AK, 4:400; PP, pp. 55–56)—cannot therefore be understood as an analysis of universal moral obligation. For it is only after they have learned to think of themselves as sensuously encumbered pure intelligences that Kant’s students will view their sense of duty in the required way—as the subjective surfacing of a pure inner law—rather than (for example) as the outcome of imposed civil obligations. Rather than analyzing a moral obligation to which all individuals are subject, Kant’s appeal to the sense of duty for its own sake is thus a means of subjecting certain individuals to the mode of obligation peculiar to university metaphysics as a particular moral culture or paideia.19 It is the means by which Kant induces his readers and students to relate to their moral sense, not as something that might be satisfied through the attainment of worldly ends—personal tranquillity, civil peace—but as the obscurely immediate command of a higher intelligence within them. It was just this mode of acceding to moral duty

19. Compare the more general characterization of the “mode of [ethical] subjection” in Foucault, The Use of Pleasure, p. 27.
that the civil philosophers feared would result in the subordination of external civil duties to those known through inner illumination.

We are now in a position to elucidate the true significance of Kant’s triumphant conclusion to section 1 of the *Groundwork*: “Thus, through the moral knowledge of common human reason, we have arrived at its principle which, admittedly, it does not thus think abstractly in a universal form, but which it does have always before its eyes and uses as a norm of judgment” (*AK*, 4:403; *PP*, p. 58). According to Kant, this first step into the metaphysics of morals—the step from ordinary to philosophical consciousness—is motivated, not by moral pedagogy, but by a “natural dialectic” inherent in man’s moral being:

The human being feels in himself a powerful counterweight to all the commands of duty presented to him by reason as so worthy of esteem—the counterweight of his needs and inclinations, the entire satisfaction of which he sums up under the name of happiness. . . . [And] from this arises a natural dialectic, that is, a propensity to rationalize against those strict laws of duty and to cast doubt on their validity. . . .

In this way the *common human reason* is impelled, not by some need of speculation (which never touches it so long as it is content to be mere sound reason) but on practical grounds themselves, to leave its own sphere and take a step into the field of practical philosophy. (*AK*, 4:405; *PP*, pp. 59–60)

Here of course we recognize, not a natural dialectic, but the discontent with man’s sensuous nature—that is, with all of the goods lumped under this pejorative—that has been incited through the inculcation of Kant’s metaphysical anthropology. The desire for a philosophical clarification of “ordinary” (prudential, eudaemonistic) morality is not something that simply occurs to “common human reason” but is something induced in those undergoing a certain form of self-problematization. Only those who learn to relate to themselves as beings whose pure intelligizing is threatened by their sensible inclinations come to think of themselves as beings in need of philosophical clarification in order to achieve moral perfection.

The sectarianism of Kant’s philosophy arises directly from this way of eliciting the desire for it. By requiring his students to recognize their personal moral impurity such that it can only be rectified through the purification brought by his metaphysics of morals, Kant is demanding exclusive adherence to his doctrines and school. Not the least disturbing aspect of modern Kantianism is the degree to which it follows the master’s example in this regard. Christine Korsgaard, for example, also insists that Kantian philosophy is the natural outcome of human reason: “Philosophy is ordi-
nary human reasoning rendered persistent. . . . Kant’s view, as I understand it, is that a person who starts out reasoning in some perfectly ordinary way . . . finds himself on a route that has no natural stopping place short of the unconditioned Ideas of Reason and the metaphysical perplexities to which they sometimes lead.”20 Korsgaard’s identification of Kant’s philosophy with universal reason thus leads her to treat assent to the Kantian ethic as the only path open to a rational person: “I am saying that if you are a truly rational agent, you must accept Kantian morality.”21 A philosophical school that embeds its teachings so deeply in its students’ sense of self—grounding its doctrines in their induced longing for a higher true morality—will be constitutionally predisposed to intellectual and moral sectarianism.

Section 2: Teaching Transcendence

Having secured an audience disposed to view itself as the bearer of a pure but latent moral law, in section 2 of the *Groundwork* Kant shows how this law may be revealed, requiring his students to rise from “popular moral philosophy to the metaphysics of morals.” This transition takes the form of a series of arguments designed to “deduce” the moral law and to show the necessity of metaphysics for obtaining this insight. Commentary on section 2, however, too often overlooks the fact that these arguments are conditioned by Kant’s powerful reinvocation of the culture of metaphysics in which the formal purity of principles is embedded in spiritual purity of the (transhuman) being who beholds them:

All moral concepts have their seat and origin completely a priori in reason. . . . They cannot be abstracted from any empirical, and therefore merely contingent, knowledge. In this purity of their origin is to be found their very worthiness to serve as supreme practical principles. . . . We ought never, as is permitted and even occasionally necessary in speculative philosophy, make the principles depend upon the particular nature of human reason. Since moral laws should hold for all rational beings [vernünftige Wesen] as such, we should instead derive them from the general concept of a rational being as such. In this way, we should first completely expound morality as pure philosophy, that is, as metaphysics, independent of the anthropology required for its application to man—as can be readily done in this wholly abstract [abgesonderter] type of knowledge. [AK, 4:411–12; PP, pp. 65–66]

21. Ibid., p. 65.
As before, through its grounding in the anthropology (and cosmology) of rational being, Kant’s argument for the theoretical need for a pure philosophy (metaphysics) relies on the existential need for moral purity incited by the metaphysical paideia. Those who attempt to derive moral concepts from man’s empirical nature are therefore not just philosophically mistaken but morally lax; moral purity can only be achieved through the exercise of metaphysical abstraction itself. Conversely, Kant regards the cultivation of this “wholly abstract type of knowledge” as something far more sublime than the adoption of a correct philosophical method, for the method of metaphysics itself holds the key to the purification of human souls:

We know well that without possessing such a metaphysics it is vain—I will not say to arrive at a speculative judgment of the moral element of duty in everything dutiful—but that it is impossible, even in ordinary and practical usage, particularly that of moral instruction, to ground morals on their genuine principles and thereby to create pure moral dispositions [Gesinnungen], grafting them onto human souls [Gemüthern] for the highest good of the world. [AK, 4:412; PP, pp. 65–66]

Framed in this manner, Kant’s arguments for linking the possibility of a pure moral law to the necessity of metaphysics take on a powerfully ascetic or self-transformative character, functioning as a means for grooming the “pure moral dispositions” valorized by the culture of university metaphysics.

Kant, however, presents these arguments as if they were solving a philosophical problem, namely, the problem of showing how a categorical imperative is possible. We can show the possibility of technical imperatives (“imperatives of skill”)—the rules of geometry, for example—by demonstrating their analytic necessity for achieving a particular technical end, such as the construction of a mathematical figure (AK, 4:417; PP, p. 70). Further, we can show the possibility of prudential imperatives as the empirically necessary means to certain kinds of happiness; although here human disagreement over the ends of empirical happiness, and the uncertainty of their worldly attainment, means that prudential imperatives lack the unified and unconditional character of the moral law (see AK, 4:418–19; PP, pp. 70–71). How though, asks Kant, can we show the possibility of the moral law’s categorical imperative, given that this is by definition unconditional, hence independent of all empirical ends or goods capable of showing its necessity as a means?

This is the problem, Kant argues, whose solution hinges on the transition to metaphysics, which enables the philosopher to transcend the world of empirical ends and means and to propose a “solely a priori” solution.
his conception of metaphysics, as the discipline permitting access to a domain where thinking natures act independently of all external empirical ends, Kant’s solution is to propose that the mere thought or concept of the categorical command might itself reveal its propositional content—and to this degree its possibility—and independent of all need to relate this command to some empirical object or end: “In this task we want first to inquire whether the mere concept of a categorical imperative may not also provide its formula, containing the only proposition that can be a categorical imperative” (AK, 4:420; PP, pp. 72–73). The only proposition that can be a categorical imperative is, of course, “act only in accordance with that maxim through which you can at the same time will that it become a universal law”—because a universal law is the only one capable of commanding the will through the mere thinking of its idea, independent of all sensible ends and desires.

In the light of our preceding commentary, it should already be clear that this chain of arguments is very far from what Kant claims it to be, namely, the elucidation of a moral principle already present in human moral consciousness. In grounding the categorical imperative’s theoretical necessity in an (induced) existential desire to behold its pure form, Kant’s metaphysical anthropology imbues his deduction with a distinctively self-transformative character and function. In this setting, the reader’s readiness to assent to Kant’s deduction of the categorical imperative—his preparedness to accept a formulation solely on this basis of its having been thought—is driven by the induced longing to join the pure intelligences who know and act through sheer intellection. Approaching Kant’s deduction of the categorical imperative in this way—treating it as a spiritual exercise promising access to a transcendent reality—provides a revealing insight into a procedure that many commentators have found difficult to reconcile with standard forms of deduction.22

We have already noted that Kant takes the crucial step towards showing the possibility of the categorical imperative and, with it, the necessity for metaphysics by speculating that through the mere a priori (metaphysical) thinking of its idea or form, independent of all empirical ends and experiences, it might be possible to have insight into its propositional content. This content is in fact the necessity that all subjective ends or wills be conformed to a universal law or general will. What demands our attention now

is the suddenness and speed with which Kant converts this speculation into a self-demonstrative truth:

When I think of a hypothetical imperative in general I do not know in advance what it will contain, until I am given its condition. But when I think of a categorical imperative I immediately know what it contains. For, since the imperative contains, beyond the law, only the necessity that a maxim conform to the law, while the law contains no condition to limit it, there is nothing remaining to which the maxim should conform except the universality of a law as such; and it is this conformity alone that the imperative properly asserts to be necessary.

There is therefore only a single categorical imperative and it is this: Act only in accordance with that maxim through which you can at the same time will that it should become a universal law. [AK, 4:420–21; PP, p. 73]

In ruling out the possibility of a discursive deduction of the moral law from empirical ends, Kant is simultaneously opening up the possibility of a nondiscursive intellectual intuition of the law. The notion of this intuition only makes sense inside the anthropology and cosmology of university metaphysics. The background idea is that concepts issuing directly from the divine intelligence, prior to their embodiment in spatiotemporal things, are self-declarative for a human intelligence whose purity permits it to participate in divine intellection. To the extent that it abstracts from the material things and ends through which pure concepts are diffracted and thereby rises to meet these concepts as they stream from the divine mind, such an intelligence knows their meaning and truth through immediate insight. It does not have to analyze this discursively from the scatter of appearances in space and time. The intuitional form in which Kant reveals the categorical imperative—“when I think of a categorical imperative I immediately know what it contains”—may thus be regarded as his performative personification of this purified intelligence, allowing him to claim insight into a principle lying beyond “the particular nature of human reason.” In short, Kant’s way of demonstrating the possibility of the moral law, through the sheer thinking of its concept, should be seen as the exemplary pedagogical performance of an exercise in self-transcendence. It is the spiritual charisma attaching to this exercise—the implicit claim to transcendent participation in the pure intelligizing of the moral law prior to its embod-

iment in the world of space, time, and utility—that validates the otherwise extraordinary claim that the possibility of the categorical imperative may be shown through the mere thinking of it.

In the more reflective treatment of this theme in the Critique of Practical Reason, Kant declares that our incapacity to provide a discursive deduction of the moral principle is the flipside of our extraordinary capacity to apprehend it as a “fact of reason”:

The moral law is given in the manner of a fact of pure reason of which we have a priori awareness and which is apodictically certain, even if it is granted that we can find no example in experience that follows it exactly. Hence [even though] the objective reality of the moral law can be proved through no deduction, or by any efforts of theoretical reason, whether speculative or empirically supported . . . it is nonetheless firmly grounded in itself [steht für sich selbst fest]. [AK, 5:47; PP, pp. 177–78]

It is surely remarkable that modern Kantians, purporting to offer a rational reconstruction of Kant’s insight, simply repeat his claim to behold a self-grounding, self-declarative object of intellectual intuition. Rawls, for example, after asking whether the procedure for reaching the categorical imperative is constructed, answers: “No, it is not. Rather, it is simply laid out.”24 Even more dramatically, Dieter Henrich argues that rather than being reached via empirical or logical ratiocination, the moral principle belongs to a “structure of recognition” in which it outstrips all our attempts to justify it, appearing in the form of a demand for “approval” so emphatic that someone who asks for a justification of the principle before approving “has already lost sight of it.”25

This way of regarding the moral principle, as something revealed in and to a higher self, only makes sense once we have understood Kant’s demonstration of the categorical imperative as an exercise in self-transcendence. For this demonstration is indeed a version of the long-standing Christian-Platonic spiritual exercise whereby, abstracting from merely spatiotemporal knowledge, the metaphysician activates the higher intellect he shares with God, thereby participating in the self-authenticating principles of an intellect that creates what it thinks. Doubtless it will seem odd to many that the voice of Kantian reason should sound so similar to the voice of God. But this will seem the less so the more we understand that the exercise through which Kant listens to reason is in fact a version of that through

which Christian-Platonists attuned themselves to the emanations of the
divine intellect. Despite the layers of discursive argument wrapped around
it—necessary insulation against centuries of Lutheran suspicion of self-
sacralizing Platonism—the deduction of the categorical imperative in
*Groundwork* 2 remains recognizably one of Hadot’s “acts of mental con-
centration and renunciation” leading to “transfiguration of the person-
ality.” Rather than being recovered from ordinary moral consciousness,
the categorical imperative—together with the associated principles of au-
tonomy (the autarky of the pure intellect), humanity (“rational being as
an end in itself”), and the “kingdom of ends” (communion in the *spiritus
mundi*)—must be regarded as the goals of an exercise in self-transformation
promising access to a spiritual elite.

**Section 3: Faith in Metaphysics**

Having led his students from their ordinary moral consciousness to
moral philosophy and from thence to the metaphysics of morals, in section
3 of the *Groundwork* Kant seeks to guide their final step: “from metaphysics
of morals to the critique of pure practical reason,” as the section’s title reads.
For Kant, critique means stepping beyond his primary method—the dem-
stration of metaphysical principles through the immanent clarification
of moral consciousness—in order to obtain a final reflection on the grounds
permitting such a demonstration. For us, having redescribed this demon-
stration in terms of the pedagogy of *Schulmetaphysik* and the exemplary
exercise in self-transcendence, critique will emerge as the final exercise in
self-transformation through which students are inducted into the morals
of metaphysics.

As in the *Critique of Pure Reason*, Kant’s procedure involves locating an
antinomy or aporia inside reason itself and then showing that this problem
can only be resolved by adopting the standpoint of critique. Reflecting back
on the prior two stages of the *Groundwork*, Kant thus observes that he has
provided a grounding for the moral law by invoking the idea of freedom as
rational autonomy. Yet this grounding, he now declares, does not appear to
show why anyone should take an interest in the moral law or subject them-
selves to it (see *AK*, 4:446–49; *PP*, pp. 94–97). Kant now argues that this
problem arises from the apparent circularity of the relation between the idea
of freedom and that of the moral law: “We take ourselves as free in the
order of efficient causes in order to think ourselves under moral laws in the
order of ends; and we afterwards think ourselves as subject to these laws
because we have ascribed to ourselves freedom of the will” (*AK*, 4:450; *PP*,
p. 97). As a result:
If someone asked us why the universal validity of our maxim as a law must be the limiting condition of our actions, and on what we base the worth we assign to this way of acting—a worth so great that there can be no higher interest anywhere—and asked us how it happens that a human being believes that only through this does he feel his personal worth, in comparison with which that of an agreeable or disagreeable condition is to be held as nothing, we could give him no answer. [AK, 4:449–50; PP, p. 97]

It will come as no surprise to learn that Kant’s way out of this carefully constructed problem lies nowhere else than in the metaphysical anthropology, which he is now prepared to call on explicitly. Man, says Kant, belongs to the sensible world to which he is attached by his passive sensibility and in which he knows himself and his actions only as phenomenal appearances given to the understanding. At the same time, however, he also belongs to the intelligible or noumenal world in which he participates through the spontaneous activity of his rational nature, which he must suppose is the transcendental ego underlying his empirical subjectivity. It is through this image of *homo duplex* as the nexus of the intelligible and sensible worlds that Kant claims to resolve the apparent circularity between the concepts of freedom and the moral law: “For we now see that when we think of ourselves as free we transfer ourselves into the intelligible world [Verstandeswelt] as members of it and recognize the autonomy of the will along with its consequence, morality; whereas when we think of ourselves as under obligation, we regard ourselves as belonging to the sensible world and yet to the intelligible world at the same time” (AK, 4:453; PP, p. 101). Kant thus treats the gap between man’s intelligible and sensible natures as breaking the circuit between freedom and the moral law. In giving rise to a certain inner tension, it is this gap—rather than the concept of intelligible freedom as such—that leads man to take an interest in a pure moral law and, in fact, to regard himself as bound by it:

Hence, in spite of regarding myself from one point of view as a being that belongs to the sensible world, I shall recognize that, as intelligence, I am subject to the law of the intelligible world—that is, to the reason that contains this law itself in the idea of freedom, and so to the autonomy of the will; consequently I must look on the laws of the intelligible world as imperatives for me, and on the actions conforming to this principle as duties. [AK, 4:453–54; PP, p. 100]

The key to the possibility of a metaphysics of morals, therefore, lies in the idea of a world of intelligences and our higher selves as members of it.
For, in the difference between this viewpoint and his knowledge of himself as a passive member of the sensible world, man experiences his “sensed dependency” on the intelligible world, feeling himself bound by its laws and thereby taking a metaphysical interest in morality. This explanation though, says Kant, marks the outermost limit of philosophical reflection on the possibility of a metaphysical moral law. For, while it may thus be shown that we take an interest in the moral law through the idea of the intelligible world, we are unable to know how we come to take this interest or just what the reality of the intelligible world and its freedom might be. On the one hand, Kant claims that the idea of his membership in the intelligible world is just one that naturally occurs to man: “This kind of conclusion must be drawn by a thinking man from all the things that are presented to him” (AK, 4:451–52; PP, p. 99). On the other hand, through this same unguided reflection, the thinking man spontaneously becomes aware that he may have no direct knowledge of the intelligible world, owing to the passive character of his sensibility, which confines human understanding to the domain of empirical appearances. For Kant, man’s self-awareness of his own dual nature therefore both drives his interest in the metaphysical world of spontaneously self-legislating intelligences yet ensures that this world will be the telos for a moral deportment rather than an object of metaphysical theory. To view the figure of homo duplex in this way—as a need of reason rather than as one of its objects—is to adopt the critical attitude towards it.

In the light of our redescription of the Groundwork, however, it will already be clear that Kant’s critical reflection on the interest in metaphysics is wholly internal to his metaphysical anthropology and paideia. Not only does this anthropology configure the division between man’s intelligible and sensible natures whose tension is supposed to drive the interest in the metaphysical world, it also erects the screen of sensibility designed to ensure that this world remains a matter of moral interest rather than theoretical knowledge. Modern commentary on this set of issues is preoccupied with showing that Kant’s division between the intelligible and sensible is not grounded in two ontological worlds—the noumenal and phenomenal—but in two “standpoints” that humans must take on their actions in “this world.” In her attempt to free Kant from suspicion of belief in the metaphysical reality of the intelligible world, Korsgaard thus comments that:

On what I take to be the correct interpretation, the distinction is not between two kinds of things, but between the beings of this world insofar as they are authentically active and the same beings insofar as we are passively receptive to them. The “gap” in our knowledge exists not because of the limits of experience but because of its essential na-
ture: to experience something is (in part) to be passively receptive to it, and therefore we cannot have experiences of activity as such. [C, pp. 203–4]

In failing to grasp the self-transformative function of homo duplex, however, the “two-standpoints” reading reveals itself to be nothing more than a particular execution of this function. Learning to view the two worlds as simply two standpoints—learning, that is, to take a purely practical non-ontological interest in the intelligible world—is the direct result of coming to relate to oneself as the bearer of a dual—intelligible and sensible—nature. For only someone who relates to themselves in this way will make the otherwise extraordinary statement that they can have no direct knowledge of their role as spontaneous intellectual beings due to the essentially passive character of their sense-based understanding. In other words, only someone who has been trained to believe that they are the bearers of a spontaneously active intellect lying outside their human senses will orient themselves to this level of being by declaring it to be beyond experience. That Korsgaard’s own “two-standpoints” reading subserves this particular intellectual deportment is clear from her version of this statement: “As thinkers and choosers we must regard ourselves as active beings, even though we cannot experience ourselves as active beings, and so we place ourselves among the noumena, necessarily, whenever we think and act” (C, p. 204).

As the epigraphs from Wittgenstein suggest, the significance of the statement that we cannot know the noumena lies not in what it says but in what it does to the one who says it as part of the specific “work on oneself” whose instrument it is. By declaring that the world of spontaneous intelligible beings cannot be an object of human experience, Kantians orient themselves to it, as an object of metaphysical longing and moral faith. Using the metaphysical anthropology to position the intelligible world as a reality lying beyond human understanding, Kant is able to incite the desire to participate in this world in the only way available to humans: by treating it as a moral orientation or standpoint for action in this world. This spiritual exercise is the basis of Kant’s critical reflection that while it is possible to think such noumenal ideas as that of the intelligible world, these must never be treated as objects of theoretical knowledge, being acceded to instead only for the moral transformation that they work in us:

In any case, the idea of a pure intelligible world, as a totality of intelligences to which we ourselves belong as rational beings (although on the other side we are also members of the sensible world), always remains a useful and permitted idea for the purposes of a rational faith
[vernünftigen Glaubens], even if all knowledge stops at its boundary—
useful and permitted for producing in us a lively interest in the moral
law by means of the noble ideal of a universal kingdom of ends in
themselves (rational beings), to which we can belong as members only
when we carefully conduct ourselves in accordance with the maxims of
freedom as if they were laws of nature. [AK, 4:462–63; PP, p. 108]

Like the university metaphysicians who preceded him, Kant thus uses
the metaphysical anthropology to induce belief in the world of rational be-
ings and its laws. He does so by deploying the gap between man’s higher
rational and lower sensible nature to incite the desire for a pure and puri-
fying metaphysical knowledge of morality, which can only be satisfied
through revelation of the categorical imperative. In treating the motivating
idea of membership in the intelligible world as one that just occurs to the
ordinary intelligence, Kant thus folds his account of the foundations of the
metaphysics of morals in on itself, thereby blocking further inquiry into
this idea. Far from indicating metaphysical skepticism, Kant’s critical de-
claration that the intelligible world lies beyond the reach of theoretical un-
derstanding is thus the means by which he converts this world into an object
of metaphysical faith. In short, Kant treats the idea that induces the interest
in metaphysics as an idea in which human beings are already interested,
transmitted to them via moral feeling from a world lying beyond knowl-
edge—but therefore beyond doubt—hence the object of a metaphysical
faith, admitting of no further explanation or inquiry.

Exit

We have however offered further explanation and inquiry. We have
shown that the three sections of the Groundwork cannot be properly un-
derstood in terms of the metaphysical recovery of a law binding on rational
beings from an ordinary moral consciousness in which it is already con-
tained. Rather, they are better understood as stages in the spiritual grooming
of a particular intellectual deportment—one that will regard true morality
in terms of the commands of a pure rational being acceded to through the
purifying discipline of metaphysics. Rather than eliciting the need for moral
philosophy from the “natural dialectic” between man’s intellectual and sen-
suous natures, section 1 inculcates the anthropology of homo duplex in order
to incite the desire for metaphysics, presenting this to students as the only
means of purifying their sensuous inclinations and realizing their higher
intellectual natures. So too, rather than showing that the pure idea of a cat-
egorical imperative may only be thought by a “pure practical philosophy”
or metaphysics, section 2 grounds this idea in the spiritual purity and pres-
tige of the metaphysical sage. This is the exalted personage whose charis-
matic intuition of a moral principle removed from all empirical content
and discursive interpretation takes place through an exemplary exercise in
self-transcendence. Finally, we have seen that section 3 of the *Groundwork*
does something quite other than show that man’s interest in the meta-
physics of morals arises inevitably from the gap between his intelligible and
sensible natures. Rather, this section seeks to elicit an interest in metaphysics
from readers whose guided recognition of themselves as divided beings is
designed to incite precisely this interest. As we have seen, not the least in-
teresting aspect of modern Kant commentary is that it is typically carried
out by readers who think of themselves (and everyone else) as divided in
just this way. As a result, they treat metaphysical ethics, not as an interest
cultivated by those committed (by fate or choice) to a certain kind of moral
life, but as the morally necessary pursuit of all humans insofar as they fulfill
their vocation as rational beings. This quasi-confessional character of the
Kantian moral culture—its commitment to a single true path to moral sub-
jecthood—helps to explain its aggressive self-certainty and its inclination
to sectarianism.