

Kant's Questions: What is the Human Being?

Patrick R. Frierson

June, 2009

*The present version of this book is a draft. Please do not quote without permission.
Comments are greatly appreciated and may be sent to frierspr@whitman.edu*

Chapter 1: Kant's Transcendental Anthropology

“As to the subject matter with which we are concerned, we ask that people think of it . . . as the foundation of human . . . dignity. Each individual equally, then, may reflect on it himself . . . [Our work] claims nothing . . . beyond what is mortal.”

— Francis Bacon, *New Organon*, quoted by Kant in the *Critique of Pure Reason* (Bii)

“I myself am a researcher by inclination. I feel the entire thirst for cognition and the eager restlessness to proceed further in it, as well as the satisfaction at every acquisition . . . [but] I would feel by far less useful than the common laborer if I did not believe that this consideration could impart a value to all others in order to establish the rights of humanity.”

— Immanuel Kant, from private notes written in 1764-5 (20:44)

In one of his lectures, Kant is recorded as having laid out his view of philosophy as a whole:

The field of philosophy . . . can be reduced to the following questions: What can I know? What ought I to do? What may I hope? What is the human being?

Metaphysics answers the first question, morals the second, religion the third, and anthropology the fourth. Fundamentally, however, we could reckon all of this as anthropology. (9: 25, cf. 11:249)

The term “anthropology” may seem odd here for contemporary readers. We are accustomed to thinking of “anthropology” as a specific academic discipline that studies variations between people in different cultures. Kant, by contrast, uses the term anthropology in its original sense, as the study (-logy) of human beings (anthropos). Thus Kantian anthropology includes comparisons between different people at different times, but it also includes – and even focuses on – general features of human beings as such. Anthropology, then, is simply the discipline that answers the question, “What is the Human Being?” That is how the term will be used throughout this book.

That said, for anyone familiar with Kant’s philosophical works as a whole, the claim that “we could reckon all of [philosophy] as anthropology” might be confusing. At the end of his life, Kant published a book entitled *Anthropology from a Pragmatic Point of View*, but this book could hardly be said to include Kant’s most important contributions to the questions of human knowledge, obligation, and hope. In fact, Kant’s *Anthropology* is striking for being deeply empirical, while Kant’s most developed answers to the questions of knowledge, obligation, and hope emphasize that these questions must be answered non-empirically, or a priori. In his *Groundwork*, Kant even goes so far as to emphasize a distinction between “pure moral philosophy,” which most fundamentally addresses the question “What ought I do?” and “moral anthropology,” which is secondary and merely adds empirical details. Kant’s *Anthropology from a Pragmatic Point of View* relates to this secondary, empirical aspect of human knowledge, obligation, and hope, and does not articulate the most important dimensions of Kant’s answers to the three questions that, supposedly, can all be “reckon[ed] . . . as anthropology.”

In this chapter, I show that Kant did develop what I will call a “transcendental anthropology” that answers all the major questions to which Kant reduces philosophy. The term “transcendental anthropology” is taken from Kant’s handwritten notes, in which he refers to an “anthropologia transcendentalis” (RA 903, 15:395). But what is transcendental anthropology? In his Critique of Pure Reason, Kant defines “transcendental *philosophy*” as “the system of all principles of pure reason” and includes “an exhaustive analysis of all human cognition a priori” (A13/B27).¹ Kant includes the metaphysics and epistemology of the Critique of Pure Reason as part of such a transcendental philosophy, but in later works also refers to the a priori parts of his moral philosophy (4:390; 5:113; 6:272; 8:381) and even his aesthetics (5:266, 270) as “transcendental.” Kant’s “Critical” philosophy (named for his Critiques of Pure Reason, Practical Reason, and Judgment) thus lays out what we might call his transcendental anthropology, a system of principles of human beings that is a priori and thereby “elevated . . . out of empirical psychology” (5:266).

At this point, one might be inclined to read “transcendental” as having the same meaning as “transcendent,” and thus taking transcendental anthropology to somehow refer to that aspect of human beings that transcends ordinary experience, or our animal nature, or something of that sort. In the same way that God might be seen as ultimately transcendent, we might want to study the transcendent aspect of human beings, through art, perhaps, or by talking about our immortal souls. Kant, however, sharply contrasts his transcendental philosophy from traditional philosophies of the “transcendent.” For Kant, “transcendental anthropology” is a kind of “self-knowledge of the understanding and of reason” (RA 903, 15:395). By this he does not mean simply that in knowing human beings, we know ourselves, since this would be true for empirical investigations of human beings as well. Instead, he refers here to the fact that in transcendental anthropology, one knows oneself from-within rather than looking at one’s psychology from the stance of an observer.² In a sense, this is the most *immanent* sort of self-knowledge, and hence sharply contrasted with the sort of transcendence of God. “From-within” here does not imply that transcendental anthropology is “introspective” in any traditional sense. One way of making this distinction clear can be seen in the case of choosing a course of action. Observing human beings, one might say that what a person chooses in a particular case will be determined by accidental aspects of her environment of which the person is only barely conscious. Or one might introspect and say that one’s behavior in a particular instance was caused by, say, a combination of

¹ For the variety of ways in which Kant uses the term “transcendental,” see Howard Caygill, *Kant Dictionary*, pp. 399-404.

² Often the distinction here is described in terms of a difference between “first person” and “third person” perspectives. This way of describing the distinction is helpful, as long as one does not mistakenly think of “introspective” states as first person. The first person point of view is the point of view *from which* one decides what to think or do, not a point of view *on* oneself. The third person point of view is a point of view *on* something, whether oneself or something else. (The second person point of view, which we will briefly discuss in chapter 11, is the perspective of a “you,” that is, one who is *addressed* by another.)

anger and exhaustion. As we will see in the next chapter, Kant's empirical anthropology focuses on these sorts of causal explanations of behavior. But when one is actually making a choice, one doesn't consider these accidental and unconscious influences as the bases for one's choice. One looks for various reasons for action, and even if these bases include what one might in another context see as mere causes of action (say, one's desiring something), they have a different character when one considers them to be reasons to act; they serve not as explanations for behavior but as justifications for it. From-within the context of deliberation, one's anger appears not as a necessary cause of action, but as a candidate reason for acting, a reason that one may either endorse or reject. Throughout his transcendental anthropology, Kant is interested in offering accounts of what is involved from-within in the processes of thinking, judging, choice, and appreciation of beauty.

As should be clear even from my brief account of action, the from-within perspective of transcendental anthropology involves an important evaluative or normative dimension. When explaining behavior non-transcendentally, one looks at what the causes of action are, and one need not evaluate whether these causes are "good." Even the question of whether, say, anger is a good cause of an action seems misguided. But when thinking about behavior (or judgments, or choices) transcendently, one looks at the reasons for that behavior (or . . .), and these reasons invite evaluation. Anger might have caused the behavior, but we can still ask whether it was a good reason for doing what one did. And from within the perspective of deliberation, this is just the sort of question that one asks, not merely when deciding what to do, but also when deciding what to believe, or how to judge about something, or even whether something is beautiful. The normative question – "Is this a good reason for people to do/think/feel such-and-such?" – arises within the context of transcendental anthropology.

Along with this from-within, normative perspective on human beings, Kant's transcendental anthropology employs a distinctive style of argument. "Transcendental" arguments in Kant proceed from some "given" to the conditions of possibility of that given. Thus Kant's Critique of Pure Reason is an extended argument exploring the conditions of possibility of experience. From the perspective of oneself as an experiencer of the world, one can think about what must be the case in order for one's experience to be possible, and Kant argues that in order for humans to have the kind of experience that we have, the world must contain substances, laws of causality, and other features, and human cognition of it must be limited in various ways. Similarly, the Critique of Practical Reason argues from the validity of the moral law that we find from within the perspective of moral deliberation and evaluation to various conditions of possibility of that validity.

In sum, Kant's *transcendental* anthropology focuses on what can be known about human beings *a priori* through an examination of their basic mental faculties "*from-*

within” that specifically attends to the *conditions of possibility* of *normative* constraints on human beings.

To see how this transcendental anthropology gives rise to Kant’s three Critiques of Pure Reason, Practical Reason, and Judgment, we need to say just a bit more about the specifics of Kant’s conception of the human being. Within both his empirical and his transcendental anthropology, Kant argues for a threefold division of human mental states into those of cognition (thoughts), volition (desires and choices), and affection (feelings). Each different aspect of human beings is governed by its own a priori principles that are prescribed by a distinct higher cognitive power (5:196). In the *Critique of Judgment*, looking back on his philosophy as a whole, Kant uses a chart to show how his entire transcendental philosophy can be understood as a transcendental anthropology of these different human faculties (5:198).³

Core aspect of the human being	Cognitive power that prescribes principles for it	A priori principles	Application to	Relevant Critique	Relevant Question
Cognition	Understanding	Lawfulness	Nature	<i>Critique of Pure Reason</i> (1781/1787)	What can I know?
Feeling	Judgment	Purposiveness	Art (and teleology)	<i>Critique of Judgment</i> (1788)	What may I hope?
Desire/Volition	Reason	Final End	Freedom	<i>Critique of Practical Reason</i> (1790)	What ought I do?

³ The first four columns are taken directly from Kant's own work, though I've edited them and modified terminology a bit. I've added the final two columns to show the connection with Kant's writings and his central questions. This tidy chart is, of course, a bit too tidy. Kant's account of human judgment (and the related faculty of feeling) is not limited to art, but includes an important discussion of teleology in nature that sets the stage for his accounts of human history. And this Critique of Judgment is not the sum of what Kant has to say about “what we may hope”; Kant even explicitly claims that his Religion within the Boundaries of Mere Reason is where he “has tried to complete the third part of his plan,” that is, to answer the third question (11:429). Moreover, Kant’s discussion of the faculty of cognition in the Critique of Pure Reason is not limited to the role of the understanding. And while the notions of “nature” and “freedom” fit well to some extent with the distinction between cognition (or the understanding), the Critique of which involves laying out the fundamental categories that structure all of nature, and desire (or practical reason), the Critique of which lays out the fundamental law of human freedom, the fit is not exact. Human cognition is also free in an important sense, and human desire is governed by laws of nature in an important sense. In fact, working out the specific way in which humans exist as both free and bound by nature will turn out to be a guiding thread throughout the whole of Kant’s transcendental anthropology. Despite problems in its details, this taxonomy provides a guiding thread to discerning the way in which Kant's transcendental philosophy is a transcendental anthropology.

Insofar as human beings have three core aspects (a claim we will see Kant defend in the next chapter), a from-within interpretation of the principles that govern each of these will exhaust the transcendental answer to the question “What is the human being?” and will thereby subsume the whole of philosophy.⁴The philosophical works of Kant’s “critical period,” starting in 1781, aim to carry out just this project.

I. What can I know?

The Critique of Pure Reason as transcendental anthropology of cognition

Kant most famous and important work, the *Critique of Pure Reason*, was published in 1781 after a “Silent Decade” during which he published virtually nothing. The work was and is his magnum opus, the work that defined him as a philosopher. Its specific focus is anthropological in that it focuses on a particular human capacity: “getting to the bottom of the faculty we call the understanding and . . . the determination of the rules and boundaries of its use” (A xvi). Kant is not interested merely in the empirical question of how the understanding operates, but in giving an account of the rules under which it must operate and the limits that these rules imply for how far we should seek to extend our knowledge. In the process, Kant aims to answer the question “What can I know?” as that question applies specifically to the “objective validity” of “a priori concepts” (A xvi) that is, in the question “what and how much can the understanding and reason cognize free of all experience?” (A xvii). Kant defends a metaphysics that consists in a priori claims about the nature of the world⁵ and lays out an epistemology that limits the scope of such claims, all through a transcendental anthropology of cognition.

For Kant, metaphysics involves “a priori synthetic” claims. An *a priori* claim is one that is universal and necessary and thus not based merely on empirical generalizations. Many a priori claims, however, such as the claim that “bachelors are unmarried,” merely unpack our concepts and thus do not tell anything substantive about the world. Kant calls these sorts of empty claims “analytic” because they merely “analyze” concepts, showing what is already contained within them. Other claims, by contrast, are “synthetic” because they synthesize, or put together, concepts that, in themselves, are distinct. For example, the claim that “Hondas are reliable” or “honey is

⁴ Kant adds that this sort of transcendental anthropology functions as a needed restraint for “self-conceit and egoism” (RA 903, 15:394). Elsewhere, Kant explains that egoism has three forms: a logical egoism that inhibits proper knowing, a moral egoism that inhibits proper choice and action, and an aesthetic egoism that inhibits proper feelings (7: xxx). Thus one can read Kant’s three *Critiques* as three arenas within which Kant emphasizes the non-egoistic (universal and a priori) aspects of our from-within perspectives on knowledge, action, and feeling.

⁵ Metaphysics must be saved from the skeptical and empiricist doubts (e.g. Hume’s) that would call into question the possibility of any a priori metaphysics at all, and it needs be limited against the pretensions of those (e.g. Descartes and Leibniz) who claimed a priori knowledge that transcends the (empirical) world.

sticky” are synthetic claims; reliability and stickiness are not part of the *definitions* of Hondas and honey. Of course, those claims are also empirical, rather than a priori. But there are some claims, such as “every change has some cause,” that are both synthetic and a priori. Here one does not merely define changes as having causes (because then one could ask whether the ripening of fruit is really a “change” if there is no obvious cause) but makes a contentful claim about all changes in the world. But this claim is not based on empirically generalizing our experience that all changes have causes. We have not actually observed causes for every change; and were someone to claim that a particular change lacked a cause, we would insist that the cause had simply not yet been discovered.⁶ When we make the claim that every change has some cause we are claiming that there is a necessary connection (hence, a priori) between these two distinct concepts (hence, synthetic). The need for a metaphysics that is at once synthetic and a priori raises “the general problem” of the Critique of Pure Reason: “How are synthetic judgments possible a priori?” (B19, cf. xxx prolegomena).

Kant’s answer to this question depends upon conceiving of metaphysics as a subset of transcendental anthropology. From the beginning of his Critique, Kant makes clear how radically human-centered his metaphysics is, comparing the fundamental shift in thinking embodied in this Critique to the revolution in astronomy effected by Copernicus:

Up to now it has been assumed that all our cognition must conform to objects; but all attempts to find out something about them a priori through concepts that would extend our cognition have, on this presupposition, come to nothing. Hence let us once try whether we do not get farther with the problems of metaphysics by assuming that the objects must conform to our cognition, which would agree better with the requested possibility of an a priori cognition of them, which is to establish something about objects before they are given to us. This would be just like the first thoughts of Copernicus, who, when he did not make good progress in the explanation of the celestial motions if he assumed that the entire celestial host revolves around the observer, tried to see if he might not have greater success if he made the observer revolve and left the stars at rest. (Bxvi, my emphasis)

Kant’s reversal is as radical as Copernicus’s but moves in the opposite direction. Astronomers before Copernicus thought of the earth – and thus human beings – as the center of the universe. Copernicus’s radical shift in perspective was to think of the earth

⁶ The development of modern quantum mechanics both problematizes and confirms this claim. On the one hand, the dominant Copenhagen interpretation of quantum mechanics *does* posit uncaused changes, or at least changes that are undetermined by their causes. On the other hand, Einstein’s reaction to this interpretation was precisely to argue that its proponents had simply failed to look hard enough to find the relevant “hidden variables.” (See Albert Einstein et. al., ‘Can Quantum-Mechanical description of Physical Reality Be Considered Complete?’ *Physical Review* **47**, 777-780; and N. Bohr, ‘Can Quantum-Mechanical description of Physical Reality Be Considered Complete?’ *Physical Review* **48**, 696-702. I discuss the relationship between contemporary physics and Kant’s transcendental anthropology of cognition in chapters eight and ten.)

on which human beings reside as just one planet among many, all circling around the sun. Copernicus thus moves humans out of the center of the universe.

By contrast, Kant moves human cognition into the center of metaphysics. He begins by isolating an assumption of prior metaphysics, the assumption that in order to know anything about the world, our judgments about the world have to conform to the way the world really is. Kant claims that this assumption has made progress in metaphysics impossible and thereby proposes an alternative to all previous philosophical attempts at metaphysics. Broadly speaking, previous philosophers – especially during the 17th and 18th centuries⁷ – were either “rationalists” who sought philosophical systems based upon reason alone or “empiricists” who sought the ultimate foundations of knowledge in experience. But empiricists fail to account for the aprioricity of metaphysics, while rationalists fail to properly account for its synthetic status (by mistakenly overestimating what reason alone can do), and Kant’s Copernican turn is based on the thought that empiricists and rationalists fail because both are looking for a way to make human cognitions fit onto an independently given world of objects. Kant sees that there is better hope of showing the possibility of a priori synthetic judgments if one assumes instead that the world of objects must conform to the structure of human cognition.

Let us once try whether we do not get farther with the problems of metaphysics by assuming that the objects must conform to our cognition, which would agree better with the requested possibility of a priori cognition of them, which is to establish something about objects before they are given to us. (B xvi)

By assuming that the world must conform to our cognition, it is possible to have knowledge of the world that is based on the structure of our cognition rather than by induction from what we experience in that world. By itself, this claim does not get beyond merely analytic claims. The world must conform to my cognition in that bachelors in the world must be unmarried, but this is not a substantive claim about the world. To get such claims, our cognition must justify synthetic claims that are knowable a priori because they structure the way the world has to be.

Kant’s next claim about human nature both limits the scope of this Copernican turn and helps show how it functions to make substantive (or “synthetic”) a priori knowledge possible. Kant claims that human cognition has both a passive and an active component.

Our cognition arises from two fundamental sources in the mind, the first of which is the reception of representations (the receptivity of impressions), the second the faculty for cognizing an object by means of these representations (spontaneity of concepts); through the former an object is given to us, through the latter it is thought in relation to that representation . . . If we call the receptivity of our mind

⁷ In his “History of Pure Reason,” Kant actually traces this distinction between empiricists and rationalists back to Aristotle and Plato (see A854/B882), but it is most commonly associated with the major empiricists (Locke, Berkeley, and Hume) and rationalists (Descartes, Spinoza, and Leibniz) of the early modern period.

to receive representations insofar as it is affected in some way sensibility, then . . . the faculty for bringing forth representations itself, or the spontaneity of cognition, is the understanding. It comes along with our nature that intuition [that through which cognition relates immediately to an object] can never be other than sensible . . . The faculty for thinking of objects of sensible intuition, on the contrary, is the understanding . . . Thoughts without content are empty, intuitions without concepts are blind. (A50-1/B74-5, cf. A19-20/B33-4)

Despite the deluge of technical terminology, Kant's point here is simply that when human beings think about objects, these thoughts have two components, an active component by virtue of which we think about the objects, and a passive component by virtue of which our thoughts are thoughts of objects. When one merely thinks about relations between concepts, one makes use of an active capacity for thought, but this capacity is not directed towards any real objects. It is, as Kant says, "empty." And when one merely "takes in" the world without actually conceptualizing it, one does not even really see what one is exposed to; one's taking-in of the world is effectively "blind." Knowledge of a real world involves receiving "impressions" from the world and processing them using one's concepts.

Kant's appeal to sensibility, or to "intuitions"⁸ given through sensibility, limits the scope of the Copernican turn. Kant does not claim, and need not claim, that everything about the empirical world is determined by the structure of human cognition. Because we have a receptive faculty, human beings have a lot of empirical knowledge that we learn *from* the world, such as the knowledge that there are mountains in the Pacific Northwest of the United States, that water freezes at 32 degrees Fahrenheit, that dogs and cats cannot interbreed, that large material objects are made of small molecules, etc. And there are other claims that are false, but if true, would have to be discovered empirically, such as the existence of the Loch Ness monster, or fairies, or solid crystalline spheres rotating in the heavens. For such empirical knowledge, cognition must conform to the world around us. The world will not have fairies in it just because we believe in fairies, nor will it cease to have molecules if we cease to (or do not yet) believe in molecules. Kant's Copernican turn justifies the possibility of some substantive a priori knowledge of the world, but it does not justify claiming that one can know everything about the empirical world simply by reflecting on one's cognitive capacities.

But Kant also argues that the distinction between intuitions and concepts (and relatedly between sensibility and the understanding) provides for the possibility of a priori knowledge of the world that goes beyond mere conceptual analysis. This is possible, for Kant, because even our receptivity to the world around us has an a priori structure to which the world must conform.

⁸ Kant's notion of an "intuition" is quite different from our contemporary use of the term to refer to judgments of which one is immediately certain. For Kant, an intuition is "that through which [a cognition] relates immediately to objects" (A19/B33). An intuition is thus something like an immediate, pre-conceptual, sensory awareness of an object.

I call that in the appearance which corresponds to sensation its matter, but that which allows the manifold of appearance to be intuited . . . I call the form of appearance. Since that within which the sensations can alone be ordered and placed in a certain form cannot itself be in turn sensation, the matter of all appearance is only given us a posteriori, but its form must all lie ready for it in the mind a priori and can therefore be considered separately from all sensation. (A20/B34)

Humans' capacity to be affected by the world has a particular structure. We perceive objects because of a human faculty of sensibility, and it should be possible to develop an a priori science of the principles of this sensibility. Moreover, precisely because sensibility is a faculty of intuitions rather than of concepts, an a priori science of sensibility will not proceed simply by unpacking concepts, and thus may provide a way of justifying claims that are both a priori and synthetic.

Kant's approach here is made clearer when he turns to the details, where he argues that space and time are the two a priori intuitions that structure all empirical intuitions. Both are a priori because we cannot think of the world as non-spatial or non-temporal and because we could never think of succession without an already-given temporal structure nor of external objects without an already-given spatial structure. Both are a priori intuitions, rather than concepts, because they are represented as given structures within which particular objects appear, rather than as constructed concepts under which objects fall.⁹ Kant further reiterates the status of space and time as a priori intuitions by pointing out that these intuitions underlie the success of geometry (in the case of space) and arithmetic (in the case of time), both of which give us synthetic a priori knowledge.¹⁰ In geometry, for example, one does not *measure* shapes to discern their properties empirically, but one also does not merely analyze the concepts of those shapes. Instead, one must make use of mental, spatial images (or axioms about mental space) to connect concepts that are not already connected by definition. Thus, for example, we can know (without empirically measuring every triangle in the world) that the sum of the interior angles of any triangle makes two right angles, even though this fact is not contained in the definition of a triangle. As an intuition, space is able to ground synthetic claims. As an a priori intuition, it grounds these claims a priori. "Thus," Kant says, "our explanation alone makes the possibility of geometry as a synthetic a priori cognition comprehensible" (A25/B41).

The understanding, like sensibility, has an a priori structure, and the heart of Kant's Critique of Pure Reason focuses on showing that the a priori structure of the understanding is a necessary condition for the possibility of experience. Kant's argument for this claim is the most difficult part of the Critique of Pure Reason, and Kant revised the argument extensively between the first and second editions of the

⁹ For these arguments, see A22-5/B37-40 (space) and A30-2/B46-8 (time). The precise details of these arguments are controversial and the success of the arguments is contested. For discussion, see Allison 1990, Guyer xxx.

¹⁰ A25/B40-1.

book.¹¹ The essence of Kant’s argument is twofold. First, he looks to “logical function of the understanding in judgments” for “the clue to the discovery of all pure concepts of the understanding” (A70/B95). The basic idea here is that once we are committed to seeing experience as a product of the operation of the understanding, one can look to the way that the understanding is used in the construction and relating of judgments to discern the fundamental categories of understanding that are used in the construction and relation of objects within experience. Since the operation of the understanding with respect to judgments is the provenance of a relatively straightforward logic, one can simply use what Kant calls “general logic” as a clue to the transcendental anthropology of the understanding. For example, Kant argues that because all relationships within judgments must be either categorical (x is y), hypothetical (if x, then y), or disjunctive (x or y), the understanding must have corresponding a priori categories for cognizing objects. Respectively, these are: inherence and subsistence (cognizing objects as things with properties), cause and effect, and community (objects as reciprocally interacting). Overall, Kant lays out a famous (or infamous) table of twelve basic categories of the understanding, drawn from a similar table of twelve functions of thinking.

Of Quantity

Unity
Plurality
Totality

Of Quality

Reality
Negation
Limitation

Of Relation

Of Inherence and Subsistence (*substantia et accidentis*)
Of Causality and Dependence (cause and effect)
Of Community (reciprocity between agent and patient)

Of Modality

Possibility – Impossibility
Existence – Non-existence
Necessity – Contingency

¹¹ For very good detailed studies of Kant’s key arguments here, see Allison, Ameriks, Guyer, Henrich, and Longuenesse.

This table proposes only what the content of a priori categories of the understanding would be, but Kant still needs to show that experience is the sort of thing that requires any such categories. In particular, given that objects are presented in terms of spatial and temporal structures of sensibility, one might wonder whether any further contribution from the understanding is needed. For that purpose, Kant lays out a detailed argument for the necessary role of the understanding in making cognition of objects possible. The key move in this argument is Kant's claim that experience is possible only by virtue of a twofold "unity" on the side of both the "object" and the "apperception" of that object. Kant takes the term "apperception" from Leibniz, for whom it referred to something like explicit consciousness of an object as opposed to mere unconsciousness awareness of it, the difference between hearing sounds when one is asleep and listening to them when one is awake. For Kant, "apperception" refers to "the 'I think' [that] must be able to accompany all of my representations, for otherwise something would be represented in me that could not be thought at all" (B131-2). This section of the first *Critique* is among the most complicated and controversial arguments in all of Kant's writings, but Kant's basic idea is to connect experience of unified objects with a unified "I think" and thereby with the categories, as the principles of any such unity. First, Kant argues that in order for one to unite different representations together into consciousness of a single object, those different representations must be held together in a single consciousness. To have a cognition of a purple cow, it will not do for one person to have a representation of the color purple and another to have a representation of a cow. These different representations, to be united into a cognition of a single object, must be united by a single consciousness.

[C]ognitions . . . consist in the determinate relation of a given representation to an object. But an object is that in the concept of which the manifold of a given intuition is *united*. Now, however, all unification of representations requires unity of consciousness. Consequently the unity of consciousness is that which alone constitutes the relation of representations to an object. (B 137)

The relevant unity of consciousness here is not that "subjective unity of consciousness which is a determination of inner sense" (B139), that is, it is not an introspective awareness of oneself as united throughout time. Rather, the relevant unity is the from-within unity by virtue of which one becomes conscious of and makes justifiable claims about objects in the world.

Now just as the table of judgment provided Kant's "clue" to discovering a set of a priori categories, so he argues here that we see, in the form of judgment itself, a necessary appeal to this transcendental unity: "a judgment is nothing other than the way to bring given cognitions to the objective unity of apperception. That is the aim of the copula *is* in them . . . For this word designates the relation of the representations to the original apperception and its [apperception's] *necessary unity*, even if the judgment itself is empirical" (B 141-2). In other words, in making a judgment such as "mangos are delicious," one not only makes an empirical claim about the world but implicitly asserts

the necessary unity of the I that holds together “mangos” and “delicious” in a single “I think.” And now Kant comes to his punchline:

Therefore the manifold [of different representations], insofar as it is given in *one* empirical intuition, is *determined* in regard to one of the logical functions for judgment, by means of which . . . it is brought to a consciousness in general. But . . . the categories are nothing other than these very functions for judging, insofar as the manifold of a given intuition is determined with regard to them. Thus the manifold in a given intuition also necessarily stands under categories. (B 143)
Thus if, e.g., I make the empirical intuition of a house into perception through apperception of its manifold, my ground is the *necessary unity* of space . . . This same synthetic unity, however, if I abstract from the form of space, has its seat in the understanding, . . . in the category of *quantity*. (B162)

Our “apperception” of objects is unified through organizing various representations together by means of a priori concepts of the understanding such as unity, reality, or causation. Even spatial and temporal properties of an object only become *properties* of an *object* by subsumed together by the understanding, under concepts such as “unity” or “property (of a substance).” These a priori concepts, or categories, provide both a from-within structure by means of which a human knower can connect representations in such a way that those representations remain part of a single “I think,” and they provide the framework according to which a mere “manifold” can become a coherent *object* of experience.

At this point, one clarificatory warning is needed lest one think that Kant’s transcendental anthropology does (or claims) more than it really does. Rene Descartes famously emphasizes the importance of the “I think,” claiming that he could be absolutely certain of the claim, “I think therefore I am” and reasoning, even further, that the nature of the “I think” implies the simplicity, unity, and ultimately immortality of the human soul. For Kant, however, such inferences mistakenly treat the transcendental unity of apperception that makes (empirical) cognition possible as itself a possible *object* of (such) cognition. In his transcendental deduction, Kant emphasizes that while the “transcendental synthesis of the manifold” makes me conscious “*that I am*,” it reveals neither what “I am in myself” nor how “I appear to myself” (B157). Kant diagnoses as “Paralogisms” (false arguments) attempts to get from the formal requirements of the “I think” to substantive claims about such an I (see A341/B399-A404/B431). For example, the inference from the necessity of a unified “subject” of thought to a unified “substance” that thinks erroneously takes the formal category of “subject” to have objective meaning, when “pure categories (and among them also the category of substance) have in themselves no objective significance . . . unless an intuition is subsumed under them” (A348-9), but one can have no intuition of the “I think” that unifies all intuitions. The only objective cognition that one can have of oneself, then, is of “our own subject only as appearance” and this is available “through inner sense” (B156). And Kant argues in a famous “Refutation of [Cartesian] Idealism” (B274-9) that knowledge of this empirical self is actually secondary to knowledge of

external objections, since “the determination of my existence in time is possible only by means of . . . actual [persisting] things that I perceive outside myself” (B 275). Kant’s appeal to the transcendental unity of apperception is an explanation of cognition from-within, one that shows that objective empirical cognition requires unifying one’s representations by means of a priori categories. It thereby shows that the objective world must be unified in that way, but it makes no objective claims about the apperception that unified that world.

Having laid out the a priori structure of both sensibility and the understanding, Kant turns to the way in which these two different cognitive faculties work together to structure the world of experience. By showing how humans’ a priori categories work with sensibility to structure the empirical world, Kant’s “system of all principles of pure understanding” provides the a priori metaphysics promised in his Preface. The specific details of the various ways in which these faculties combine is both complicated and contested, but one example (Kant’s most famous) is sufficient to give a sense for his general strategy here. One of the principles that Kant defends as a principle by which human beings structure the objective world is that of cause and effect: “All alterations occur in accordance with the law of the connection of cause and effect” (B232). Kant’s argument proceeds by considering what is necessary in order for a set of perceptions to be considered perceptions of alteration (or, more generally, of something happening). Kant distinguishes between mere subjective perceptions and objective experience. To have objective experience, we must organize perceptions in accordance with categories. But to have experience of objective alteration (succession), perceptions must be ordered in accordance with the category of cause—effect. If ordered using inherence—subsistence (seeing some perceptions as properties of others) or part—whole (seeing each perception as part of a whole), the sequence of one’s perceptions would not refer to an objective sequence, since objectively, one supposes that the properties of the thing exist at the same time as the substance and one supposes all the parts of the thing to exist at the same time.¹² Given the Copernican turn, to say that one must order perceptions in a certain way is just to say that the objects of those perceptions must in

¹² Kant gives, as an example of a purely subjective sequence, the perception of a house, starting with the roof, then the windows, then the door, and then the chimney, but one doesn’t suppose that objectively speaking there really is first a roof, then windows, then a door, etc. On the other hand, one might actually suppose that the order of one’s perceptions does correspond to an objective order. For this, Kant gives the example of a boat. One perceives a boat upstream, a boat midstream, and a boat downstream, and one supposes not that these are different parts of a complicated stream-wide boat, but that in reality – that is, objectively – the boat is moving. And Kant then considers what sort of concepts one would have to impose on one’s set of perceptions to order them in such a way that one considers their order objective. His answer is that the perceptions would have to be thought of as though they have to occur in the order in which they do. And this necessary sequence of perceptions must be according to some rule. But necessary sequence according to a rule is just what one thinks of when one thinks of the relation between cause and effect. So if one is to think of the order of perceptions order as referring to an objective order, one must impose the concepts of cause and effect on those perceptions.

fact be ordered in that way. By imposing an aspect of the structure of the human understanding – the category of cause and effect – on the subjective flow of perceptions in inner sense, human beings are able to perceive the world as (and thereby structure the world as) a series of causally determined changes.

Throughout his proofs – for the necessity of space, time, the categories, and causation – Kant does not provide merely empirical claims about human cognition. Consistent with his insistence on transcendental anthropology, Kant looks at cognition from-within, arguing that certain cognitive presuppositions are necessary conditions of the possibility of justifying the claims that we make about the world. Because, from-within, we take a priori mathematics to be justified, we must assume that space and time structure our world. Because we can make justified empirical claims about objects, we must be organizing and unifying the diffuse manifold of intuition into coherent cognition. And because some of this cognition is of objective succession, we must apply the categories of cause and effect to structure the world of our experience. In the end, then, Kant’s Critique of Pure Reason provides a transcendental analysis of the human faculties of sensibility and the understanding that elucidates their a priori structure and the contributions of this structure to experience of an objective world. The human beings for whom experience – and a priori synthetic judgments – is possible are finite beings dependent upon sensibility and also spontaneous free thinkers. Moreover, given Kant’s Copernican turn, this transcendental anthropology provides both an epistemology that delimits what we can know and a metaphysics, in that the world itself must conform to the structures of human cognition. Metaphysics and epistemology turn out, in Kant’s hands, to be reckoned as (transcendental) anthropology.¹³

With his analysis of the way in which sensibility and the understanding combine to structure a knowable empirical world and his defense of several specific a priori principles of human cognition to which that empirical world must conform, Kant completes the first part of his answer to the question “What can I know?” But Kant’s

¹³ The radicality of Kant’s position here can be seen in one of the most famous criticisms of it. Bertrand Russell critiques Kant’s account of a priori knowledge on the grounds that, among other things, Kant’s relies on contingent facts about human nature. As Russell explains,

The thing to be accounted for is our certainty that the facts must always conform to logic and arithmetic. To say that logic and arithmetic are contributed by us does not account for this. Our nature is as much a fact of the existing world as anything, and there can be no certainty that it will remain constant. It might happen, if Kant is right, that to-morrow our nature would so change as to make two and two become five. This possibility seems never to have occurred to him. (xxxxref)

Russell’s problem arises because Kant seems to ground the a priori necessity of truths of mathematics (and even logic) on the conditions of possibility of our sensing and thinking about the world. Because we (human beings) perceive the world in Euclidian space and time and think about it using various logical categories, truths of geometry, arithmetic, and logic must be true in any world that is to be a world for us. But, Russell suggests, that makes these truths ultimately contingent upon human nature, and a truth that is contingent upon human nature cannot be the sort of a priori – and hence necessary – truth that Kant sought in his Critique of Pure Reason. (See Frierson 2009: xxx-xxx; Guyer 2006: 66-67, 81-2; and vanCleve xxx:xxx-xxx for Kantian responses to Russell’s objection.)

transcendental anthropology of cognition involves two further elements as well. One of these is not continued until a subsequent work. In his *Metaphysical Foundations of Natural Science*, Kant argues that “natural science presupposes . . . metaphysics of nature,” which includes not only the “laws that make possible the concept of a nature in general,” laid out in his *Critique of Pure Reason*, but also laws that “concern [themselves] with a particular nature of this or that kind of things, for which an empirical concept is given, but still in such a manner that, outside of what lies in this concept, no other empirical principle is used for [their] cognition” (4:469-70). The general idea is that the nature of the human mind is such that if it cognizes, say, material bodies, then it will have to cognize them in particular ways. Given the Copernican turn, these necessary ways of cognizing would be a priori principles of material bodies themselves. What this implies, for Kant, is that the basic principles of physics itself can be seen as a sort of transcendental anthropology of cognition. In his *Metaphysical Foundations of Natural Science*, Kant argues that given the barest concept of matter in motion, one can derive a priori such claims as the conservation of matter (4:541), the claim that all motion is relative to a particular frame of reference (4:487), and Newton’s laws (e.g. that “every change in matter has an external cause” (4:543) and “in all communication of motion, action and reaction are always equal” (4:554)). For Kant, not only the most basic metaphysical claims about the universe, but even the most basic physical laws of that universe are a priori conditions of the possibility of experiencing nature at all. And they are a priori conditions precisely because they reflect transcendental structures of human cognition. For Kant, Newtonian physics is transcendental anthropology.

The second further element of Kant’s transcendental anthropology of human cognition comes in the second part of the *Critique of Pure Reason*. There Kant turns from human sensibility and understanding to human “reason.” The a priori forms of sensibility and the understanding are “constitutive” of experience; the objective world that we experience is constituted by conformity with these forms. Reason, by contrast, presents ideals that regulate humans’ pursuit of knowledge, but that are not constitutive of that knowledge itself. In the context of this *Critique* reason can be understood as that most philosophical of cognitive powers, the one that constantly asks the question, “Why?” And in asking this question, reason constantly seeks the “unconditioned,” that is, an answer that does not itself require a further explanation. For Kant, this search for the unconditioned plays itself out in every area of knowledge. Reason prompts humans to seek the causes of phenomena in our world, and then the causes of these causes, and so on. It prompts us to look for the constituent parts of the objects in our world, and the constituent parts of those, and so on. It prompts us to develop an idea of an *ens realissimum*, a most real being whose perfection is not conditioned by anything. In all of these ways, reason drives human knowers to discover more and more about the world in which we live.

But with this regulative function of reason comes a dangerous illusion, one that “does not cease even though it is uncovered and its nullity is clearly seen into by

transcendental criticism” (A297/B353).¹⁴ Reason drives humans to learn more and more about their world in a search for the unconditioned, but this impulse naturally generates the illusion that there is in fact an unconditioned to be found. Thus as we search for the smallest and most basic particle, or the first cause of the universe, or the most perfect possible being, reason incites us to assume that there is in fact a most basic particle, a first cause, or a perfect being. The second half of Kant’s Critique of Pure Reason is focused on showing why these assumptions are unwarranted assumptions. The basic argument in each case is that the conditions of possibility of objective cognition conflict with the nature of the unconditioned, such that these ideal goals of reason are the sorts of things that could never exist in a world structured by human forms of intuition and understanding. While we should still make use of reasons’ ideals as ideals guiding inquiry into the objective world, we cannot make sense of them as really existing in that world.

The details of these arguments are unnecessary in this brief account of Kant’s transcendental anthropology, but it is worth looking in detail at one of Kant’s specific discussions: the third “antinomy of pure reason.”¹⁵ In his antinomies, Kant shows that when one assumes that the ideals of reason must apply to the objects of possible experience, it is possible to prove contradictory positions on many of the most important questions of metaphysics. The third antinomy starts by proving both that “Causality in accordance with the laws of nature is not the only one from which all the appearances of the world can be derived. It is also necessary to assume another causality through freedom” (A444/B472) and that “There is no freedom, but everything in the world happens solely in accordance with laws of nature” (A445/B473). The proof of freedom depends upon the claim that no explanation of any appearance can ever be sufficient if it is conditioned by a further explanation. This proof thus appeals to reason’s need to seek the unconditioned. The argument against freedom draws on Kant’s argument for the necessity of causation in the second analogy (discussed above) to show that there could never be a cause that was not itself the effect of a further cause, since any act of causation would have to be an event in the world, and every event must have its own cause in turn.

Kant’s “solution” to the antinomy involves two components, both of which are important for understanding Kant’s transcendental anthropology. The first important aspect of Kant’s solution is to draw attention to the merely regulative function of the ideals of reason. Insofar as the demand for sufficiency is a regulative ideal, the empirical

¹⁴ As he put it at the very start of his Critique,

Human reason has the peculiar fate in one species of its cognitions that it is burdened with questions which it cannot dismiss, since they are given to it as problems by the very nature of reason itself, but which it also cannot answer, since they transcend every capacity of human reason. (Avii)

¹⁵ Kant does not introduce this discussion in the context of a discussion of human freedom. Instead, the issue of freedom comes up in the context of thinking about whether, in general, “causality in accordance with the laws of nature is . . . the only one from which all appearances . . . can be derived” (A444.B472), but Kant quickly shifts to talking about the possibility of human freedom as a causality that grounds appearances in the world.

investigation of natural causes should constantly seek a “sufficient” cause for any appearance by investigating causes as far back as they will go. But any cause that one in fact finds will, by virtue of being itself an appearance, call for investigation into its own cause. Thus we must seek sufficiency in our explanations, but we will never find it.

But Kant’s resolution of the third antinomy introduces a second aspect of his transcendental philosophy which is crucial to understanding Kant’s answer to the question “What is the Human Being?” So far, my discussion of Kant’s transcendental anthropology of cognition has focused on the positive contribution that this anthropology can make towards a robust metaphysics of nature: given that our empirical knowledge depends upon the structure of our sensibility and understanding, we can prove important claims about the empirical world a priori. But Kant points out that this positive contribution entails “a very strange result . . . , namely, that with this faculty [of cognition] we can never go beyond the boundaries of possible experience” (Bxx). We can establish a priori claims about possible objects of experience, but we cannot provide any theoretical justification for any claims at all about things that would be unexperienceable. Kant calls such things “things-in-themselves” or “noumena” and distinguishes them from objects of possible experience, which he calls phenomena. And throughout his transcendental account of cognition, Kant reminds his readers that the nature of human cognition determines only the way in which objects (of possible experience) must be, not the way in which “things-in-themselves” must be. Thus, for example, when he proves a priori that the world must be spatiotemporal, Kant claims,

Our expositions accordingly teach the reality (i.e. objective validity) of space [and time] in regard to everything that can come before us externally as an object, but at the same time the ideality of space in regard to things when they are considered in themselves through reason, i.e., without taking account of the constitution of our sensibility. (A28/B44)

For something to be an empirical object for us, it must be presented to us through our sensibility. And since the structure of our sensibility is spatial, all empirical objects must be spatial. But precisely because the spatiality of empirical objects is due to our sensibility, we are not justified in saying that such objects are spatial apart from human sensibility. The result is that Kant’s metaphysics commits him to what he calls an “empirical realism” but a “transcendental idealism.” His metaphysics is empirically real because its claims (e.g. about causality) are necessarily true of the empirical world. But it is transcendently ideal because such claims are limited to the empirical world and say nothing about what “things-in-themselves” – apart from human sensibility – might be like.

When Kant turns to the third antinomy, this transcendental idealism does significant work. Consistent with the insights of his second analogy, Kant insists that any objective alteration must be the result of causes in accordance with natural laws. But he turns in his resolution to the third antinomy to the question of “whether it is a correct disjunction that every effect in the world must arise either from nature or freedom, or whether instead both, each in a different relation, might be able to take

place simultaneously” (A336/B564). Given Kant’s transcendental idealism, the law-governed causality of the empirical world does not preclude a different kind of causality – freedom – operating at the level of things-in-themselves. In his resolution, Kant lays out this possibility in detail:

[F]or a subject of the world of sense we would have first an empirical character, through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural laws . . . [and] second . . . and intelligible character, through which it is indeed the cause of those actions as appearances, but which does not stand under any conditions of sensibility [including universal causation] and is not itself [an empirical object]. (A539/B567)

This distinction between empirical and intelligible character makes it possible for Kant to defend the possibility of what he (misleadingly)¹⁶ calls “transcendental freedom,” a power “of beginning a state from itself, the causality of which does not in turn stand under another cause determining it in . . . accordance with the law of nature” (A533/B561). We therefore cannot rule out the possibility that humans, as things-in-themselves, have an intelligible character that is transcendently free in this sense. But this intelligible character can itself be the ground of an empirical character, and one who observes this empirical character will be able to trace empirical causes for any particular action.

The result is an initially shocking but ultimately quite plausible account of the relationship between freedom and natural necessity, one that distinguishes Kant’s account from the dominant accounts of freedom and causal necessity both in his day and our own. Many philosophers are “compatibilists,” who argue that freedom is compatible with causal determination. Generally, compatibilists define freedom as determination by internal, psychological causes rather than external ones, such that if an action follows from my choice, it is free, even if my choice was determined by external factors. Other philosophers defend incompatibilism, the view that freedom and natural determination conflict with one another. Such philosophers can be either “hard determinists” who believe that every event in the world, including every human choice and action, is causally determined by some set of prior conditions¹⁷ and therefore argue that there is no room for any meaningful kind of freedom, or “libertarians” who believe that (some) events in the world are determined by human choices that these choices are not in any sense determined by prior conditions.¹⁸ Kant’s own position has aptly been

¹⁶ I call this misleading because the freedom that Kant here calls “*transcendental*” is really “*transcendent*.” It is not a condition of the possibility of experience nor evident from-within theoretical reasoning, but “transcends” any possible experience.

¹⁷ This is the dominant form of hard determinism among secular philosophers today. Other classic forms of determinism (or “predestination”) claim that every event is determined by Fate or by God.

¹⁸ Among libertarians, some (e.g. Kane) emphasize natural indeterminacy according to which certain choices are undetermined natural events and others (e.g. xxx) emphasize “agent causation,” where events that are undetermined by prior conditions are explained as being the effects of agents who are responsible for them.

called a “compatibility of compatibilism and [libertarian] incompatibilism.”¹⁹ Like incompatibilist libertarians, Kant defines freedom in a way that excludes prior causal determination of one’s choices, but like compatibilists, Kant believes that there is a way in which one can assert both that something is freely caused and that something is the result of prior empirical causes. What makes Kant unique among contemporary theories is that he preserves a thoroughgoing causal necessity but at the same time an undetermined freedom. Kant’s transcendental idealism allows him to see free things-in-themselves as grounds of the empirical world, while his empirical realism allows him to insist that within that empirical world, causation universally proceeds according to natural laws.²⁰

It is important to note that in the Critique of Pure Reason, where Kant develops this account of freedom, he does not posit that such freedom is actual. Kant’s Critique of Pure Reason is focused on what can be known a priori about the objective world, the world as it exists for human knowers. And what can be known about that world is that every event, including every human action, is causally determined by prior conditions in accordance with natural laws. Kant’s discussion of freedom and causation is directed not only against those who argue for the impossibility of freedom (say, on the basis of Newtonian physics), but also against those who would claim that one can know that freedom exists in the world. For Kant in the Critique of Pure Reason, the “single thing we could accomplish and . . . our sole concern” is “to show that . . . nature at least does not conflict with causality through freedom” (A558/B586).

The theory of freedom that Kant lays out in the first Critique is thus presented as an option that metaphysics can neither establish nor rule out. And this sort of modesty shows an important positive aspect of Kant’s limitations of metaphysics. In his Preface to the Critique, after noting how his account of cognition precludes metaphysical proofs about things like God, humans’ immortal souls, or freedom, Kant insists that while

a critique that limits . . . is, to be sure, to that extent negative, . . . because it simultaneously removes an obstacle that limits or even threatens to wipe out the practical use of reason, this critique is also in fact of positive and very important

¹⁹ Wood 1984: 74.

²⁰ There are two dominant ways that Kantians interpret this position. So-called “two-world” theorists read Kant as positing two metaphysically distinct “worlds,” a noumenal world of things-in-themselves and a phenomenal world of appearances. The former includes humans insofar as we are free, the latter humans insofar as we are determined. And the former is the “ground” of the latter. Alternatively, so-called “two-standpoint” theorists claim that Kant posits only a single world that can be thought of in two different ways, as the sum of objects of possible experience or as a merely thinkable abstraction. When thinking of the world in the former way, freedom is precluded, but not when thinking of it in the latter way. Because morality requires thinking of ourselves as free (as we will see in the next section), the “merely thinkable” perspective gets content as a practical perspective from which we hold ourselves responsible. Thus insofar as human beings take an agent-standpoint on the world, we must view human beings as free. Insofar as we take a scientific-observer standpoint, we must see everything (including human beings) as causally determined.

utility, as soon as we have convinced ourselves that there is an absolutely necessary practical use of pure reason (the moral use), in which reason unavoidably extends itself beyond the boundaries of sensibility. (Bxxv)

Or, as he puts it much more succinctly later, “I had to deny knowledge in order to make room for [practical] faith” (Bxxx). Given his Copernican turn, Kant was able to use a transcendental anthropology of cognition to justify not only epistemic claims about the nature of human knowledge but even metaphysical claims about the nature of the empirical world. But precisely because such claims are limited to objects of possible experience, Kant makes room for non-empirical claims, if there is any non-cognitive access that human beings have to things-in-themselves. And Kant finds in morality a role for reason in governing a non-cognitive sphere of human life: our volition, or “faculty of desire.”

II. What ought I do?

Kant’s moral philosophy as transcendental anthropology of volition

From the question “What can I know?” Kant turns to the question, “What ought I do?” While Kant’s transcendental analysis of cognition focused on human beings as human beings as free but finite knowers, Kant aims here to think about human beings as free but finite doers (or agents).²¹ But as in the case of cognition, Kant focuses on what human action looks like “from-within,” and in particular explores both the norms that govern such action and the conditions of possibility of being governed by those norms. Through laying out both the nature of action-guiding norms and the conditions of possibility of being bound by these, Kant is able to offer insight – though not “knowledge” in the strict sense – into what human beings are in themselves. In particular, Kant’s moral philosophy completes the argument for human freedom by showing that such freedom is not only possible, but actual²², and by laying out “laws of freedom” that govern free human beings (G 4:387).²³

²¹ Kant’s Groundwork might seem specifically to avoid developing ethics as a subset of anthropology. Kant insists in the Preface to the Groundwork “a law, if it is to hold morally, . . . must . . . hold not only for human beings, as if other rational beings did not have to heed it . . . [T]herefore the ground of obligation here must not be sought in the nature of the human being . . . but a priori simply in concepts of pure reason” (4:389). And Kant is, in fact, deeply opposed to thinking of morality as a subset of human biology or psychology, explicitly rejecting approaches to ethics that start with “conditions of human volition . . . drawn from psychology” (4:390-1). Thus Groundwork discounts what Kant calls “practical” or “moral anthropology” as merely a subsidiary part of ethics, one that “would deal only with the subjective conditions in human nature that hinder people or help them in fulfilling the laws of a metaphysics of morals” (G4:388, MS 6:217). The core of morality, Kant insists, must be “pure.” But this dismissal of anthropology at the core of morals is really only a dismissal of empirical anthropology at that core. As Groundwork progresses, the centrality of the human being as a free but finite chooser emerges clearly.

²² Both this “possibility” and “actuality” need to be interpreted carefully. Strictly speaking, the *Critique of Pure Reason* did not show that human freedom was really possible, only that the fact that human beings as they appear in the world are governed by natural laws does not necessarily preclude the possibility of

Kant's argument for the actuality of freedom is based on the nature of moral obligation. For Kant, the from-within standpoint of volition – where one seeks to discern what one ought to do – has two important features relevant to human freedom. First, anyone who asks, in the broadest sense, what to do, “must regard itself as the author of its principles independently of alien influences; consequently, . . . as the will of a rational being it must be regarded of itself as free” (4:448). All choice happens “under the idea of freedom” (4:448) because the “power of choice . . . cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim*” (6:24). This “Incorporation Thesis” claims that, from-within the standpoint of deliberation, all inclinations and incentives appear only as *candidate* reasons for action; one must “incorporate” them into one’s plans for action before they actually motivate. From-within, one sees this incorporation as something “free.”

For some contemporary Kantians, this analysis of the deliberative perspective from-within is sufficient to establish human freedom, but Kant worries that this argument does “not prove freedom as something *real*” but only as a necessary but possibly illusory “presupposition” (4:448-9). Thus Kant turns from the generic perspective of deliberation to the more specific stance of one asking the question, “what *ought* I do?”, where “ought” is specifically *moral*. In the *Critique of Practical Reason*, Kant insists that the moral “ought” is ever-present within human practical deliberation: “the

humans being free “noumenally” or “in themselves.” And, as we will see, the “actuality” that is established in Kant's moral philosophy is, metaphysically speaking, not an actuality in the objective world (since the “objective” world is a world of appearances), and epistemically speaking, not an actuality of which we can have “knowledge” in Kant's strict sense.

²³ Whereas Kant's Critical treatment of the question “What can I know?” is concentrated in his Critique of Pure Reason, his answer to “What should I do?” is scattered throughout several different texts. As early as 1765, Kant claimed to be working on a “Metaphysical First Principles of Practical Philosophy” (10:56) and in 1768 he assured a former student (Herder) that his “Metaphysics of Morals” would be done within a year (10:74). Later, he wrote that when he finished his “critique of pure reason” he would then develop his “metaphysics of nature and metaphysics of morals” (10:145). A Metaphysics of Morals would, in fact, have made a nice complement to the Metaphysical Foundations of Natural Science, and Kant's Critique of Pure Reason even suggests such a parallel (A841/B869). In fact, however, Kant's first Critical work addressing the question “What should I do?” was neither a Critique nor the promised Metaphysics of Morals but a mere Groundwork for the Metaphysics of Morals (1785). In this work, Kant aims at “nothing more than the search for and establishment of the supreme principle of morality” (4:392). This work thus addresses the question “What should I do?” only in the most general terms. Kant would subsequently write a Critique of Practical Reason (1788), which recapitulates the overall claims of the Groundwork but greatly elaborates the psychology and metaphysical implications that Kant sees as connected with this general moral theory. Not until shortly before his death would Kant finally publish his long promised Metaphysics of Morals (1797), in which he takes the general moral framework of the Groundwork and articulates a whole theory of the most important political and ethical obligations of human beings. Throughout this section, I aim to offer a synchronic account of Kant's moral philosophy drawing from all of his “Critical” works (starting with the *Groundwork*). In several cases, Kant's positions on important issues in moral philosophy shifted between these works, most notably with respect to his arguments for freedom and morality. For discussion of some of these changes, see Ameriks xxx, Sussman xxx, and xxx.

moral law, *of which we become immediately conscious (as soon as we draw up maxims of the will for ourselves)*, . . . offers itself to us and . . . reason presents it as a determining ground not to be outweighed by any sensible conditions and indeed quite independent of them” (5:29-30, emphasis added). In the process of devising and considering principles to act on, we become aware that we are bound by a moral law, one that commands obedience regardless of and independent of any other incentives. Kant considers our consciousness of the moral law a brute “fact of reason.” We are not aware of the moral law through reasoning from any “antecedent data of reason,” but rather the fact that we are under the moral law is something that “forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition” (5:31). It is something of which we are “immediately conscious” (5:29), rather than something we infer from prior knowledge.

From this fact of reason, Kant aims to establish that human beings are (transcendentally) free by showing that “a [transcendentally] free will and a will under moral laws are one and the same” (4:447, cf. 4:450, 5:28-9).²⁴²⁵ In order to establish this mutual implication, Kant draws on “common rational moral cognition” to “search for and establishment of the supreme principle of morality” (4:392). What could the supreme principle of morality be? To answer this question, Kant focuses on two (related) features of the moral ought, its independence from inclination and its universality. Moral reasons are distinguished from other sorts of reasons in that they are not tied to things that one happens to find oneself wanting. When one decides that one “should” buy gasoline for one’s car, one does so only because one thinks that such an activity will be conducive to ends that one happens to have. One can always decide to forgo those ends, and then one need not buy gasoline. But when one decides that one “ought” to refrain from falsely accusing an innocent adversary or “ought” to help a stranger in immediate pressing need, one does not see these decisions as optional in the same way. It does not matter whether the false accusation fits with other goals that one has, nor whether one cares about the stranger. *Moral* obligations do not depend upon our inclinations. Kant puts this point in terms of a distinction between what he calls

²⁴ Given this “circle” (4:450) or reciprocal implication (5:29), establishing either that one is under moral laws or that one is free is sufficient to establish both. Although Kant initially (in the *Groundwork*) tries to establish the validity of the moral law by first offering an independent and quasi-theoretical argument for human freedom, his mature position is that we must first know the validity of the moral law, which is given as a fact of reason and cannot be deduced from any prior knowledge, and only from our knowledge of this fact of reason can we know that we are free.

²⁵ The fact that we are obligated by the moral law shows us that we are free if there is a “reciprocal” relationship between the idea of a free will and the idea of a will under moral laws. Recall our discussion in section I of this chapter about Kant’s distinction between “analytic” claims and a “synthetic.” The former only includes claims about how various concepts are related to each other, unlike the latter, which includes claims about things that go beyond what is contained in the mere concept of those things. Since the claim that “a transcendentally free will and a will under moral laws are one and the same” is a claim about the relations among concepts, viz. that of a transcendentally free will and that of a will under moral laws, Kant must support this claim, often referred to as the “Reciprocity Thesis,” through an analysis of the relevant concepts. He takes up this task in the first two parts of the *Groundwork*.

“hypothetical imperatives,” which are commands that one has to obey *if* one wants to achieve some particular end, and “categorical imperatives,” which are (moral) commands that one simply *has to obey no matter what* (no “if”-clause). Relatedly, Kant argues that the moral law is *universal*: “everyone must grant that a law, if it is to hold morally . . . must carry with it absolute necessity,” going so far as to say that “the command ‘thou shalt not lie’ does not hold only for human beings, as if other rational being did not have to heed it” (4:389). Kant’s point here is not that everyone ought always act in the same way. Someone who cannot swim need not jump into a river to save a drowning child, and someone with remarkable artistic talents may have an obligation to cultivate them that others would not have. The point, rather, is that morality itself is universal, in that when one becomes immediately conscious of obligation in general, one is conscious of it as a law that binds everyone (even if it binds different people in different ways).²⁶ Another person *who is relevantly similar* to me (able to swim, or possessed of similar talents) will have the same obligations. Unlike inclinations, morality is not something that one can pick and choose. It *obligates* everyone.

Given these characteristics of morality, one might think that it would be impossible to derive a fundamental formula of morality. If all that we know about morality in general is that it can derive from neither particular inclinations nor contingent features of ourselves, then there seems to be nothing left from which to get a “principle” of morality at all. But in fact, Kant argues that the limitations on the content of the moral law actually give rise to a “formula” that encapsulates the fundamental principle of morality.

When I think of a hypothetical imperative in general I do not know beforehand what it will contain; I do not know this until I am given the condition [i.e., the end to be promoted]. But when I think of a categorical imperative I know at once what it contains. For, since the imperative contains, besides the law, only the necessity that the maxim be in conformity with the law, while the law contains no condition to which it would be limited, nothing is left with which the maxim of action is to conform but the universality of a law as such; and this conformity alone is what the imperative properly represents as 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 become a universal law. (4:420-1)

The moral law of which I am immediately conscious within deliberation is a law that commands me to act only in such a way that the bases for my actions – my “maxims” – could be bases for the actions of everyone. What is universally commanded to all is the practice of acting in a way that could be universal for all.

²⁶ This is particularly evident in those applications of the moral law that are least situation-dependent, such as “thou shalt not lie,” since in this case, the particular moral requirement is universal as well.

Kant goes on to redefine this categorical imperative based on a particular feature of human willing: human beings not only follow various practical laws, but also must act for the sake of ends (4:427). Now the moral law is not determined *by* any particular (contingent) ends, but it does determine a necessary end, “something the existence of which in itself has an absolute worth, something which as an end in itself could be a ground of determinate laws” (4:428). And Kant finds just such an end in “the human being” (4:428). This gives Kant a new way of describing the categorical imperative:²⁷ “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” (4:429).²⁸ This new formulation of the moral law puts human beings at the center of morals, not only in that the moral law is derived from a transcendental anthropology of volition, but also in that the ultimate end of morality, that which must at all times be respected, is nothing more (nor less) than the human being.²⁹

Finally, Kant adds another formulation of the categorical imperative that further enriches his transcendental anthropology and paves the way for his defense of human freedom. Kant insists that “the human being is . . . subject only to laws given by himself . . . and is bound only to act in conformity with his own will” (4:432). Kant describes this independence from external laws as “autonomy” and points out that this autonomy does not imply lawlessness, but rather that one is subject always only to one’s own laws. This may seem to be merely a recapitulation of the Incorporation Thesis, but Kant’s point here is more specific. If the moral law is to be truly universal and independent of our inclinations, then it cannot be derived from anything external to our will itself. Any *external* command would need to appeal to us for some reason, either because we feel inclined to obey it (in which case it is not truly moral) or because we *ought* to obey it (in which case its authority derives from morality, rather than grounding it). For moral laws to be truly one’s own rather than merely results of outside influences manipulating our contingent desires, autonomous lawgiving must proceed by means of laws that have no basis other than our own wills. But laws determined solely by our wills are categorical. So, for Kant, autonomy of the will [is] the supreme principle of morality” (4: 440).³⁰

²⁷ Add footnote re: controversy about whether these are really equivalent formulations.

²⁸ See too the *Metaphysics of Morals* (6:462):

Every human being has a legitimate claim to respect from his fellow human beings and is in turn bound to respect every other. Humanity itself is a dignity; for a human being cannot be used merely as a means by any human being (either by others or even by himself) but must always be used at the same time as an end.

²⁹ Precisely what Kant means by humanity here is hotly contested. See Dean, Frierson, xxx

³⁰ Kant takes this argument one step further. The moral law is “the will of every rational being as a willing giving universal law” (4: 431). But this universal sort of autonomy is tied, for Kant, to the fact that “a rational being belongs as a member to a kingdom of ends when he gives universal laws in it but is also himself subject to these laws” (4:433). Respect for humanity in oneself and others involves submitting oneself only to laws that one legislates for oneself, but also legislating laws for oneself that one also, at the

At this point, Kant has *nearly* proven that human beings are transcendently free. The principle of morality is a principle of autonomy, or *self*-governance. But to make the stronger claim that this “autonomy” is identical with transcendental *freedom*, Kant goes further. He offers an analytical, quasi-geometric proof starting with the nature of moral obligation and deriving the necessity of transcendental freedom. He poses the following problem:

Supposing that the mere lawgiving form of maxims is the only sufficient determining ground of a will: to find the constitution of a will that is determinable by it alone. (5:28)

That is, Kant considers what sort of will could be determined by a moral law that dictates only the “form” that one’s maxims must take, that is, that says only that such maxims must be universalizable, without saying anything about the “matter” of those maxims, that is, what sorts of goals one should aim for in one’s actions. Kant argues,

Since the mere form of a law . . . is not an object of the senses and consequently does not belong among appearances, . . . this form as the determining ground of the will is distinct from all determining grounds of events in nature . . . , [so] a will [determined by this ground] must be thought as altogether independent of the law of causality. (5:28-9)

From-within the standpoint of deliberation, when one considers whether or not to act on the basis of the moral law, one precisely sees this law as a law that offers nothing to one’s natural inclinations. There is, in that sense, no “natural” basis for acting in accordance with it. When one chooses to act on an ordinary inclination – say, deciding to eat an appetizing cookie – one can see oneself as “giving in” to the flow of natural causes. But because its demands are fundamentally formal, the moral law is not the sort of thing that one can merely “give in” to. It “presents it[self] as a determining ground not to be outweighed by any sensible conditions and indeed quite independent of them” (5:29-30). Thus the only will that can be truly bound by the moral law is a will that is free from sensible (that is, empirical) conditions. But freedom from determination by empirical conditions is precisely what transcendental freedom is, so a will under the moral law is a transcendently free will. Kant decries any traditional form of compatibilism as “wretched subterfuge” (5:96); “psychological or comparative” freedom, where “free” just means that “actions are caused from within,” is “nothing better than the freedom of a turnspit” (5:96-7).

To his abstract argument and violent polemics, Kant adds a more intuitive thought-experiment to show that when we reflect on actions from-within, in terms of what we take ourselves to be capable of, even apparently irresistible temptations are eminently resistible:

Suppose someone asserts of his lustful inclination that, when the desired object and the opportunity are present, it is quite irresistible to him; ask him whether, if a

same time, legislates for all rational agents. Thus human beings are not only autonomous, but autonomous members of a kingdom of ends.

gallows were erected in front of the house where he finds the opportunity and he would be hanged immediately after gratifying his lust, he would not then control the inclination. One need not conjecture very long what he would reply. But ask him whether, if his prince demanded, on pain of the same immediate execution, that he give false testimony against an honorable man whom the prince would like to destroy under a plausible pretext, he would consider it possible to overcome his love of life, however great it might be. He would perhaps not venture to assert whether he would do it or not, but he must admit without hesitation that it would be possible for him. (5:30).

The first part of this thought-experiment shows only that human beings are capable of overcoming particular sensuous desires (such as lust) when the fulfillment of these threatens more important sensuous desires (such as love of life). But the second part shows that human beings recognize in themselves an ability to overcome even our love of life for the sake of the moral law. If our love of life can motivate us to overcome our everyday sensuous desires, and our respect for the moral law can motivate us to overcome even our love of life, then there is no temptation that we are unable to overcome for the sake of the moral law.

Through examining the moral law present from-within in human volition, Kant shows that humans are transcendently free and thus “fills the vacant space” (5:49) left open by his theoretical philosophy. But Kant does more. By specifying the most fundamental principle of morality, Kant fills this vacant place “with a determinate law of . . . an intelligible world . . . , namely the moral law” (5:49). That is, Kant shows not only that human beings are free, but also that human freedom is not lawless and arbitrary but a law-governed capacity to be moral.³¹ Humans’ sense of moral obligation, properly understood, provides evidence of freedom and also gives rise to a specific principle of morality.

³¹ Just as Kant showed that the moral law implies freedom, he also shows that freedom implies morality. The point here is that if we think about what a law of freedom must be, we know that we cannot derive such a law from anything about what our natural motives and interests happen to be, since these are all determined by laws of nature. But we also know that any law of freedom, precisely because it is not based on contingent empirical details, will be the same for all rational agents. And so, for Kant, the “content” of the moral law is simply putting these formal criteria of morality into the form of a principle governing the will: whatever you do, make it the kind of thing that could be done from freedom, that is, that could be a law for anyone.

Kant's arguments for transcendental freedom are hardly beyond controversy,³² so the rest of this section focuses on two key problems that arise for the Kantian account of freedom and morality offered so far. The first problem is this: if human beings are *really free* only insofar as we submit to the moral law, Kant seems unable to account for the possibility of human beings ever being responsible for doing what is morally *wrong*.³³ If the moral law is the law of freedom, then whenever human beings act in ways that are contrary to the moral law, they must not really be free. But freedom is a condition of the possibility of moral responsibility, so whenever human beings act wrongly, they are seemingly not morally responsible for their actions. Now Kant does claim that human beings can be held responsible for acting badly,³⁴ but how can he do this?

First, as important as freedom is to his transcendental anthropology, Kant recognizes that human choosers are not *merely* morally free beings. Even from-within the perspective of human volition, we find ourselves to be *both* free beings subject to the moral law *and* members of the sensible world, subject to empirically-informed desires and inclinations. Even Kant's "pure moral philosophy" articulates an account of what morality means for beings like us, beings that participate in both an intelligible world governed by laws of freedom and a sensible world governed by laws of nature.³⁵ Because of our sensible nature, human beings have various natural inclinations that can conflict with the demands of the moral law in particular circumstances.³⁶ It is because we have such non-moral inclinations that morality takes the form, for us, of "duties" and "imperatives," commands that we *ought* to obey rather than a moral law that we simply *do* obey (4:413).

³² Among the problems with the argument that I do not discuss in detail here: Kant's appeal to a "fact of reason" makes the argument ineffective against true moral skeptics. Compatibilist freedom is, today, generally considered in much better shape than Kant's polemics allow, and even many of Kant's supporters think that his moral theory can do without a strongly metaphysical account of transcendental freedom. And Kant's account of the relation between freedom and morality poses specific problems for making sense of how finite free beings like ourselves can be bound by the moral law; if humans as empirical entities in the world are bound by natural, causal laws, how would we ever recognize the free submission of a human being to morality? In later chapters, I provide more detail about Kant's response to compatibilism (chapter nine) and the empirical expression of morally good action (chapter three).

³³ This objection has been raised in various forms by Carl Reinhold (in *Letters on the Kantian Philosophy*), Henry Sidgwick (*Methods of Ethics*), and xxx (some contemporary xxx).

³⁴ We will examine this claim in detail in chapter four.

³⁵ Thus while Kant's *Groundwork* begins by discussing the "good will" in general, which belongs to God and can also belong to human agents, Kant quickly specifies the nature of this will in such a way that it applies more particularly to human wills: "we shall set before ourselves the concept of . . . a good will though under certain subjective limitations and hindrances" (4:397).

³⁶ In his *Groundwork and Critique of Practical Reason*, Kant names the comprehensive satisfaction of these interests "happiness," and he explains that human beings are always tempted by a "principle of self-love" rooted in our sensible nature, a principle that ought to be – but isn't always in fact – restricted by the moral law rooted in our free, rational nature. The scope of natural inclinations is very broad, such that even such things as a natural philanthropy that "finds an inner satisfaction in spreading joy" (4:398) is a merely natural inclination. Thus "self-love" has different forms.

In that context, Kant distinguishes between “positive” and “negative” freedom.³⁷ Negative freedom is a “property in us . . . of not being necessitated to act through any sensible determining grounds” (6:226, cf. 4:446), while positive freedom is the property of acting through a non-sensible determining ground (the moral law).³⁸ The former, negative freedom, is necessary in order to hold human beings morally *responsible*, while the latter, positive freedom, is what constitutes the full-blown autonomy of a morally *good* agent. This might seem, however, to draw too sharp a line between these two sorts of freedom and undermine moral autonomy. In particular, if negative freedom is not identical to positive freedom, it is unclear why we should see the moral law as arising from the free choice that makes us morally responsible. Why see the moral law, that is, as any less alien to our (negatively) free selves than natural causes? In response, Kant argues for an intrinsic link between negative and positive freedom. Insofar as negative freedom is a freedom from having one’s actions governed by anything external to oneself, the only way to remain free is to make one’s only law the law of freedom, which is the categorical imperative. As one commentator has put it, “by making the [categorical imperative] its principle, the free will retains the position of [freedom].”³⁹ By contrast, “the free will that puts inclination above morality sacrifices its freedom for nothing.”⁴⁰ Not only does (negative) freedom make morality possible by freeing us from inclination, but the moral law itself is identical with the internal constitutive standards of freedom. Because the moral law specifies nothing other than the condition of freedom, any choice that is based on a principle other than the moral law is a limitation, rather than a reaffirmation, of one’s own freedom. Human beings are always negatively free, in that we *need not* let our actions be determined by forces external to us, but we are not always positively free, since we often relinquish our autonomy in the face of temptations present in the natural world.

A second problem arises given that Kant’s transcendental anthropology of desire is intended not merely to lay out the conditions of possibility of moral responsibility but also to clarify precisely what, from the standpoint of deliberation, humans find themselves obligated to do. But if Kant’s moral philosophy is supposed to answer the question, “What ought I do?” the mere formula of universal law (FUL) – “*act only in accordance with that maxim through which you can at the same time will that it become a universal law*” (4:421) – seems a bit too abstract to provide any real guidance. As Hegel famously put it, Kant’s categorical imperative is an “empty formalism,” an “abstract universality, whose determination is . . . without content.”⁴¹ Precisely because this “specific principle of morality” is purely formal, it gives only the most abstract account

³⁷ This distinction is often linked with Kant’s distinction between the “will” (German *Wille*) and “choice” (German *Willkür*). The will, which Kant identifies with “practical reason” (xxx), has positive freedom. “Choice” is the human capacity by virtue of which we have negative freedom.

³⁸ (4:446-7, 5: xxx)

³⁹ Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge University Press: 1996), p.166.

⁴⁰ Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge University Press: 1996), p. 167.

⁴¹ Hegel, *Philosophy of Right*, ed. Allen Wood, Cambridge 1991, § 135, p. 162.

of what is required of human beings. In order to complete his account of the norms that ought to govern human volition, Kant must deliver a more complete framework of normative constraints on human volition. And Hegel and others have argued that this will require “bringing in material *from outside* [to] arrive at *particular* duties [because] it is impossible to make the transition to . . . particular duties . . . from the determination of duty as *absence of contradiction . . . with itself*.”⁴² For Hegel, Kant’s abstract moral law is insufficient to provide moral content from-within; we need particular and substantive duties that cannot be derived merely from Kant’s categorical imperative.

Kant has a two-fold way of dealing with Hegel’s empty formalism objection. First, even if the categorical imperative is, in itself, an empty formalism, it can still be action-guiding in an important way. Hegel suggests that the categorical imperative would only prohibit the stealing of property, for example, if one has independent bases for thinking that property rights are good. But insofar as one tests *maxims* for action, one can evaluate those maxims based on the values implicit within them, without ascribing any independent normative weight to those values. Thus the thief who acts on the maxim, “I will steal my neighbor’s car in order to have it for myself,” commits herself to the value of private property by virtue of her end (having it *for herself*), and thus her maxim conflicts with the categorical imperative. The native Hawaiians who acted on the maxim “We will take the iron nails out of the bottom of that ship in order to make spears”⁴³ did not violate the categorical imperative (at least not directly⁴⁴) because their maxim implied a commitment to the value of spears (over ships), but not any direct commitment to institutions of private property. And to this extent, at least, even if the categorical imperative is insufficient for evaluating the moral status of *actions*, it does seem to be an important way of picking out certain *maxims* that, because they require making an exception of oneself, are morally wrong.

Second, Kant’s emphasis in the *Groundwork* on pure moral philosophy is explicitly only a *foundation* for a complete “metaphysics of morals.” Just as the empirical concept of matter is needed to move from the metaphysics of the *Critique of Pure Reason* to the basic principles of physics, empirical attributes of human beings are needed to move from a general principle of morals to specific moral duties. As the particular kinds of embodied, finite agents that human beings are, we have specific talents, needs, strengths and limitations that give rise to specific duties. The normative force of these duties comes from their connection to the fundamental moral principle by virtue of which human beings are rational, free, autonomous agents. But the specific content comes from the way that we must act in order for our empirically discoverable needs and desires to be satisfied through acting on maxims that conform to that fundamental moral principle.

⁴² Hegel, *Philosophy of Right*, ed. Allen Wood, Cambridge 1991, § 135, p. 162.

⁴³ Reference.

⁴⁴ Kant does argue that a commitment to private property is implied by any external use of one’s freedom, so the native Hawaiian’s maxims imply, for Kant, an indirect commitment to private property.

The result, when Kant turns to his *Metaphysics of Morals*, is a detailed account of human obligations in the face of our finite natures. First, Kant points out that for *embodied* free beings, freedom manifests itself through actions in an empirical world. From this empirical claim, Kant arrives at his “universal principle of right”:

Any action is *right* if it can coexist with everyone’s freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone’s freedom in accordance with a universal law. If then my action of my condition generally can coexist with the freedom of everyone else in accordance with universal law, whoever hinder me in it does me *wrong*. (6:230)

In other words, interfering with another’s activity or condition is wrong, unless that other’s activity or condition is itself wrong. This general principle of non-interference involves translating Kant’s general FUL specifically in terms of human *actions* rather than maxims, and it grounds what becomes Kant’s *political theory*. Political “coercion that is opposed to [wrong actions]” is justified as a “hindrance to a hindrance of freedom” (6:231), and Kant goes on to specify specific political rights based on increasingly specific empirical descriptions of human beings, ranging from the need for private property rights (since human beings depend upon external goods to exercise their freedom) to the need for parental rights and duties (because of the vulnerability of children). The result is a detailed *Doctrine of Right*, which lays out those actions that not only are required or prohibited by the moral law but also ought to be required or prohibited by enforceable political laws.

Second, given the structure of human volition, which always acts for the sake of ends, there must be “ends that are also duties” (6:385). Because humans are dependent upon their own abilities and those of others to accomplish any ends at all, we have obligations both to promote our own abilities, or perfections (for the sake of accomplishing whatever future goals we might find ourselves to have), and to use our abilities to help others accomplish their ends (since we, like them, are dependent upon others’ assistance).⁴⁵ From the individual abilities that we can improve in our own case, Kant derives duties prohibiting such things as suicide and lying and requiring such things as self-examination and the deliberate cultivation of “powers of spirit, mind, and body” (6:444). From our obligation to make the happiness of others an end, he derives virtues such as beneficence and gratitude and vices such as arrogance, ridicule, and contempt. Kant goes on to discuss friendship as “the most intimate union of love with

⁴⁵ We do not have a “duty” to promote our own happiness since this is something “everyone already wants unavoidably” (6:386), and Kant (confusingly) argues that we do not have a duty to promote others’ “perfection” since such perfection “consists just in this: that *he himself* is able to set his end in accordance with his own concepts of duty” (6:386). This is confusing, since Kant generally uses “perfection” to refer to nonmoral as well as moral perfections, and it is not clear that, for instance, developing a better memory requires developing that memory *for oneself*. But this confusion is relatively benign; insofar as perfections are non-moral abilities to achieve ends, the obligation to promote others’ happiness includes an obligation to promote their perfections. Insofar as “perfection” is specifically moral and thus not necessarily linked to (earthly) happiness, his reason for excluding it applies.

respect” and to mention duties that arise from specific, empirically-given “differences in rank, age, sex, health, prosperity or poverty, and so forth” (4:569).

In the end, Kant’s transcendental anthropology of desire offers a detailed answer to the question “What ought I do?” and in the process further expands on the conception of human beings as free and finite beings that Kant began in the *Critique of Pure Reason*. Not only are we free and finite *doers* as well as *knowers*, but because transcendental freedom is a condition of possibility of the moral obligation under which we find ourselves from-within the standpoint of choice, we can justifiably believe that humans are transcendently free things-in-themselves, even though we can never have “knowledge” in the strict sense of that fact. Kant’s promise in the first *Critique* that he would “deny knowledge to make room for faith” (B xxx) is fulfilled in his moral philosophy. In the process, the “faith” for which he held out hope in the *Critique* is shown to be not a *blind* faith, but a solid conviction grounded in rational arguments based on the conditions of possibility of moral responsibility. Kant calls such morally-grounded beliefs “postulates” and says,

All of them proceed from the principle of morality, . . . [which] requires these necessary conditions [such as freedom] for the observance of its precept. These postulates are not theoretical dogmas but *presuppositions* having a necessarily practical reference and thus, although they do not indeed extend speculative cognition, they give objective reality to the ideas of speculative reason . . . (by means of their reference to what is practical) and justify holding concepts even the possibility of which it could not otherwise presume to affirm. (5:132)

Kant even claims that this morally-grounded faith has a sort of “primacy” (5:119) over the knowledge that one has through empirical cognition; even one’s theoretical reasoning about the world must “accept” the postulates “as soon as these propositions belong inseparably to the practical interest of pure reason” (5:121). The transcendental anthropology of desire fills in the otherwise “vacant space” for freedom left by human cognition (5:49, 103).

Before closing this section, it is worth attending to one further, radical aspect of Kant’s transcendental anthropology of desire. In the *Critique of Pure Reason*, Kant had highlighted *three* traditional problems of metaphysics that would be stricken from the realm of knowledge – “God, freedom, and immortality” (Bxxx, A3/B7) – and in his transcendental anthropology of volition, Kant comes back not only to freedom but to the issues of God and immortality as well. As in the case of freedom (though to a different degree), Kant argues that belief in God and immortality are practically necessary. Neither God nor immortality are conditions of the possibility of moral responsibility *per se*, but Kant claims that as finite, free agents, human beings always act for some end. When Kant considers what the end of a virtuous agent must be, he argues that while the “supreme end” will be virtue alone, the “complete” end – that end from which nothing good is absent – must include both virtue and “happiness distributed in . . . proportion to morality” (5: 110). The reason for this is not that the

agent merely *wants* happiness, but that “an impartial reason” could not deliberately choose a world within which some beings “need happiness, [and are] worthy of it, and yet [do] not participate in it” (5:110). Insofar as virtuous agents seek this highest good, they must believe in whatever is necessary in order for their activity to reasonably be held to contribute to this highest good. For Kant, immortality is necessary because virtue can never be fully realized in one’s finite life but only in endless progress (5:122). God is necessary in order to ensure that happiness is doled out in proportion to virtue (5:124ff.). Only by believing in both God and immortality can our efforts towards virtue be reasonably taken to be efforts towards the highest good.

Kant’s arguments for God and immortality are much more complicated than I have suggested here, and their validity is widely disputed. For the purposes of understanding Kant’s conception of human beings, the details of these arguments are less important than the overall implication of Kant’s approach. Just as Kant in the *Metaphysics of Natural Science* makes Newtonian physics a subset of a transcendental anthropology of cognition, he here makes traditional *theology* a subset of a transcendental anthropology of volition. By the end of his transcendental anthropology of cognition, Kant had shown that the a priori structure of human cognition establishes (among other things) our ability to know a priori an empirical world as consisting of substances in causal relationships with one another, and he offers a priori foundations for natural science. Having added a transcendental anthropology of volition, Kant has laid out the a priori laws governing the realm of free human agents, and he has even established the existence of God as a necessary part of a philosophical *anthropology*.

III. What may I hope?

The Critique of Judgment as transcendental anthropology of feeling

Given the results of the previous two sections, it might seem as though Kant’s transcendental anthropology is complete. Human beings are free, finite knowers and doers, governed within each realm by a priori laws that we give ourselves. We exist as both fully free things-in-ourselves and finite, embodied appearances in the empirical world. Within the empirical world, we see ourselves and everything else as governed by natural laws. As free, we are governed by moral laws. Kant’s first and second questions – about knowledge and obligation – have been answered, and the question “What may I hope?” seems answered by Kant’s practical postulates of God and immortality. For a long time, Kant took his transcendental anthropologies of cognition and desire to complete his transcendental philosophy as whole and specifically “thought it impossible” to find a priori principles for the faculty of feeling (A21).⁴⁶

Nonetheless, shortly after finishing his *Critique of Practical Reason*, Kant set to work on a third *Critique*, which would eventually become the *Critique of Judgment* and

⁴⁶ Moreover, Kant’s *Critique of Pure Reason* seemed to include an analysis of “the logical functions of the understanding in judgments” (A70/B95) and a “transcendental doctrine of the power of judgment” (A137/B176), so a specific critique of judgment seemed out of place.

would provide the a priori laws of feeling that could complete his transcendental anthropology. By this time, Kant had made three realizations that required a rethinking of the nature of his anthropology. First, Kant came to see that his earlier rejection of feeling as a faculty capable of transcendental investigation was mistaken. Kant had rightly seen that the pleasures that humans take in what is merely “agreeable” – food, sex, reputation, baseball – are empirically rooted and thus incapable of a priori investigation. But as he continued to teach and study aesthetics, he came to see that aesthetic judgments that something is beautiful or sublime are at once subjective because rooted in feeling and taken to be universal and normative; to claim that something is beautiful is to claim that *all* others *should* find it beautiful. Thus there was a normative and potentially a priori realm of pure feeling. Second, Kant recognized that his account of the cognition of nature was incomplete in its application to the empirical world. His *Critique of Pure Reason* ensured that the world would conform to certain general structures of human cognition, but it provided no assurance that humans would be able to expand the scope of their knowledge in any systematic way. Finally, Kant’s moral philosophy was incomplete in its application to the empirical world. The *Critique of Practical Reason* provided an a priori argument to show that the end human beings are obligated to promote – the highest good – is possible, but it provided no basis for this possibility in the observable order of nature. Kant’s attempt to “deny knowledge in order to make room for belief” was insufficient to explain *how* nature and freedom relate to each other. He needed a *Critique of Judgment* to provide a “mediating concept between the concepts of nature and the concept of freedom, which makes possible the transition from the purely theoretical to the purely practical, from lawfulness in accordance with the former to the final end in accordance with the latter” (5:196, cf. 5:176). Thus Kant came to see his accounts of cognition and of volition as insufficient in their application to the empirical world, and he lost his early pessimism about the possibility of a transcendental (a priori, normative, from-within) analysis of feeling. These three realizations led Kant to complete his transcendental anthropology with a *Critique of Judgment* that would investigate the faculty of feeling through the power of judgment.

The general structure of the *Critique of Judgment* can seem perplexing, since it is divided into two halves that seem, at first blush, unrelated.⁴⁷ The first half – a “Critique of Aesthetic Judgment” – explores conditions of possibility of making justified aesthetic judgments about beauty or sublimity. The second half – a “Critique of Teleological Judgment” – lays out Kant’s philosophy of biology, within which Kant argues that for the study of living things, one must make use of teleological principles in addition to the laws of mechanical causation defended in his *Critique of Pure Reason* and *Metaphysics Foundations of Natural Science*. Moreover, while the “Critique of Aesthetic Judgment” focuses on laying out an a priori principle that governs the faculty of feeling, there is no direct reference to feeling in the “Critique of Teleological Judgment.” And while the

⁴⁷ For a defense of the claim that they are not in fact related in any significant way, see J.D MacFarland, “The Bogus Unityxxx” (1976, in Allison). For a detailed reading of the *Critique of Judgment* as a coherent whole, see Zuckert xxxx.

“Critique of Teleological Judgment,” especially with its discussion of the ultimate and final ends of nature, provides a transcendently-grounded framework for answering the question “What may I hope?,” the analysis of the beautiful and sublime seems irrelevant to answering that question. All of this can make it seem that however helpful this book might be in other respects, it cannot provide a unified transcendental anthropology of feeling that would complete Kant’s philosophy.

In fact, however, the book as a whole is unified by the principle of the purposiveness of nature, an a priori principle of judgment that provides a basis for universal norms governing feeling. Purposiveness emerges as an a priori principle of judgment in the context of Kant’s transcendental anthropology of aesthetic feeling, or taste: “It is . . . properly only in taste, and especially with regard to objects in nature, in which alone the power of judgment reveals itself as a faculty that has its own special principle and thereby makes a well-founded claim to a place in the general critique of the higher faculties of cognition” (20:244). But this a priori principle can be applied more widely than *merely* to govern aesthetic feeling:

once the capacity of the power of judgment to institute a priori principles for itself is granted, then it is also necessary to determine the scope of this capacity, and for this completeness in critique it is required that its aesthetic faculty be recognized as contained in one faculty together with the teleological and as resting on the same principle, for the teleological judgment about things in nature also belongs, just as much as the aesthetic, to the reflecting power of judgment. (20:244)

Moreover, once purposiveness is established as an a priori principle, Kant makes use of it to address the other two concerns that led Kant to posit another *Critique* in the first place: the insufficiency of his transcendental accounts of cognition and volition when it comes to applying them to the empirical world. What starts in a transcendental anthropology of (aesthetic) feeling becomes the unifying principle of Kant’s transcendental anthropology as a whole and the basis for answering the final question of Kant’s philosophy: “What may I hope?”

Before unpacking the details of this account, we should address the question of whether a transcendental anthropology of feeling is even appropriate. Recall that *transcendental* anthropology has at least three distinctive features: it is a priori, investigates humans from-within, and emphasizes normative constraints. Now human *feelings* seem ill-suited to any of these sorts of analyses. Of all aspects of human life, feelings seem to be the most empirically contingent. And even though we can introspectively examine our feelings, there does not seem to be the sort of “from-within” relationship to feeling that we have with cognition and desire. Whereas we actively think and choose, feelings seem to be things that just happen to us. This point is tied to the last; Kant’s “from-within” perspective is fundamentally normative, not a matter of how things seem to us but a matter of how we govern ourselves in thought or choice. And normativity does not seem appropriate to feeling; it is at least a bit odd to say that a person *felt* wrongly. And even if there is *some* sort of normativity governing feelings, surely what one ought

to feel cannot be derived *a priori*; can it? It might seem as though any anthropology of feeling should be empirical, a matter of introspection into what one does feel in various circumstances and observation of the feelings of others, with perhaps some prudentially-normative guidelines based on what makes for a good human life or some moral restrictions ultimately traceable to the moral principle at the heart of Kant's transcendental anthropology of volition.

In fact, Kant raises many of these concerns himself. He points out, for example, that while there are "empirically knowable" connections between objects and natural feelings of pleasure that give rise to desires for those objects, such connections are "not grounded in any principle *a priori*" and thus do not provide suitable material for a transcendental anthropology of feeling (20:206, cf. A21). (Kant calls the objects of these pleasures "agreeable.") Other objects might give rise to pleasure because they are useful in some way, and one takes pleasure in their suitability to some end. Such objects please because they are "good-for" something and their pleasure will be based in empirically knowable connections between those objects and the ends for which they are good. Of course, there is a feeling of respect for the moral law, which is both a feeling and required *a priori*, but it is required only by virtue of its connection with the faculty of desire/volition. Thus the necessity of the feeling of respect does not require a "special . . . critique of the feeling of pleasure and displeasure" but can be subsumed under a transcendental anthropology of volition (as Kant does in the *Critique of Practical Reason*). In fact, pleasure in *both* the agreeable *and* the good (whether useful or moral) can be explained by reference to the faculty of desire (or volition). Agreeable objects are the goals of hypothetical imperatives; they are objects that we find ourselves desiring, so we feel gratification in getting them (5:207). Useful objects are the necessary or helpful means to some given ends, so they provide enjoyment "only as a means" (5:207). And the morally good is the object of the categorical imperative; again, we feel gratification in the morally good because of its connection to volition.

But Kant suggests that some pleasures are caused neither by agreeable nor by good objects. These pleasures, for Kant, are judgments of "taste" or of "aesthetic pleasure" and have for their objects things that are "beautiful" or "sublime." Kant's primary focus is on the beautiful, and he structures his transcendental analysis of beauty around several key claims about how pleasure in the beautiful presents itself to us from-within: it is *disinterested* (5:204-211), it is non-conceptually *universal* and *necessary* (5:211-219, 5:235-40), and it presents its object as *purposive without a purpose* (5:219-235). For these sorts of pleasures, Kant argues, an *a priori* principle is both needed and available.

Kant's first claim – that aesthetic pleasure is disinterested – merely emphasizes that beautiful objects are neither agreeable nor good and thus cause pleasure without connection to "interest" (that is, without volition). To judge "whether something is beautiful" we consider it from-within the standpoint of "mere contemplation," and from-within that standpoint, "one only wants to know whether the mere representation of the object is accompanied by satisfaction in me, however indifferent I might be with regard to the existence of the object" (5:204-5). For example, if I want to know whether

“the palace that I see before me [is] beautiful,” I need know only whether the mere contemplation of it brings pleasure, even if, “were I to find myself on an uninhabited island . . . and could conjure up such a magnificent structure through my mere wish, I would not even take the trouble of doing so” (5:204-5, cf. 29:877-8).

The second claim brings up central dilemma that drives Kant’s analysis, the “reason why judgments of taste are subject to a critique with regard to their possibility” (5:191). The dilemma is that aesthetic judgments involve “not an empirical concept but rather a feeling of pleasure . . . which . . . is nevertheless to be expected of everyone” (5:191). For Kant, when one judges that some object is “beautiful” or “sublime” (the two kinds of aesthetic judgments), one does not ascribe to that object anything that can be expressed in general concepts; one claims only that the object gives rise to (a certain kind of) pleasure. But nonetheless takes one’s judgment to be “correct,” that is, one takes it that the object *should* give rise to that pleasure. And this is not merely a subjective claim; one takes it that the object should give rise to pleasure *in anyone*. Moreover, like other key claims in Kant’s transcendental anthropology, this insistence that the object cause pleasure in human beings generally is not an empirical-psychological claim; one does not claim that all others *will* or *do* feel this pleasure, but rather than that they *should*. Aesthetic judgments thus present a sort of normativity that is reducible neither to epistemic norms (since epistemic norms apply to the formation and application of concepts) nor moral-practical (both because aesthetics is disinterested and because practical norms require an appeal to concepts). Because aesthetic judgments, seen from-within, are governed by a sort of normativity that is irreducible to epistemic and moral normativity, a separate transcendental analysis of the conditions of possibility of this normativity is required.

The final key claim about pleasure in the beautiful provides Kant’s solution to his central dilemma, but it is also the most confusing of Kant’s claims about the beautiful. Beautiful objects incite pleasure because they are “purposive without a purpose.” What does that mean? And why would the purposiveness of beautiful objects be a ground for universal pleasure? It is in answering these questions that Kant most explicitly connects his transcendental anthropology of *feeling* with a critique of the power of *judgment*. The normative universality of aesthetic feeling is explicable in terms of purposiveness as an a priori principle of judgment that governs both experiences of beauty and our investigation of nature. By showing the connection between aesthetic feelings and purposiveness as a principle of judgment, Kant also solves the problems of incompleteness in the first and second *Critiques*.

Before showing how purposiveness explains the normativity of aesthetic judgment, Kant explains how it functions in the formation of empirical cognitions and thereby deals with an incompleteness in the transcendental anthropology of cognition offered in the *Critique of Pure Reason*. In section one, we saw how Kant argues for metaphysical principles such as universal causation, but in the *Critique of Judgment*, Kant points out that the a priori principles that make empirical cognition possible as such do not ensure the further “unity of experience . . . as a system in accordance with empirical

laws” (5:183) because they do not preclude the possibility that each different experience of succession is governed by a different causal law.

[F]or nature in general (as the object of a possible experience) that law [of causation] is cognized as absolutely necessary . . . [H]owever, the objects of empirical cognition are still . . . , as far as we can judge a priori, determinable in so many ways . . . that specifically distinct natures . . . can be causes in infinitely different many ways . . . Thus we must think of there being in nature, with regard to its merely empirical laws, a possibility of infinitely manifold empirical laws . . . ; and with regard to them we judge the unity of experience (as a system in accordance with empirical laws) as contingent. (5:183)

The first *Critique* showed that changes in the world must happen according to causal laws, but it failed to show that the set of causal laws governing the world is finite, much less that these laws fit into anything like a systematic whole within which diverse particular laws are explicable in terms of more general laws. The first *Critique*, in other words, gave no indication of the possibility of anything even approximating the grand unified theories that scientists seek. But Kant goes on,

But since such a unity must still necessarily be presupposed and assumed . . . , the power of judgment must thus assume it as an a priori principle for its own use that what is contingent for human insight in the particular (empirical) laws of nature nevertheless contains a lawful unity. (5:183)

Human cognition involves seeing the natural world in terms of objects that change over time in accordance with necessitating natural laws, but we also seek systematic interconnections amongst these objects and the laws that govern them. While it would be consistent with the conditions of possibility of experience in general for each change to be governed by its own causal law, such that the world as a whole was a “manifold” (i.e., crazy mess) of different laws, human beings cannot actually think that this is the case. Our principles for investigating the world assume uniformity that, strictly speaking, we are not justified in assuming. Kant refers to “pronouncements of metaphysical wisdom” that are “scattered about in the course of science” such as that “Nature takes the shortest path” or “the great multiplicity of its empirical laws is nevertheless unity under a few principles” (5:182, cf. 5:185). Without such cognitive rules of thumb, we could never get anywhere in terms of a systematic empirical science; we would be left with the abstract metaphysical foundations of science laid out in Kant’s earlier transcendental philosophy.

For Kant, the “power of judgment” provides the transcendental basis for these scientific rules of thumb. In general, “the power of judgment is the faculty of thinking of the particular as contained under the universal” (5:179), and it has two basic forms: *determining* and *reflecting*. Determining judgment is the power to subsume a particular under an already-given universal. When one sees a particular tree as a tree, one *determines* that the particular present object falls under one’s concept of a tree. Reflecting judgment is the power by which we are able to “ascend from the particular . . . to the universal” (5:180). When one *reflects on* the bones of an animal that is somewhat like a

human being but not quite the same, one can gradually form the concept of a Neanderthal or an Australopithecus. And whereas determining judgment has for its “rule” the (empirical) concepts that it applies,⁴⁸ reflecting judgment “requires a principle that it cannot borrow from experience, precisely because it is supposed to ground the unity of all empirical principles under equally empirical but higher principles” (5:180). The rules of thumb used in science are common-sense versions of this a priori (transcendental) principle of reflective judgment.

This transcendental principle of reflective judgment is purposiveness. “Nature specifies its universal laws in accordance with the principle of purposiveness for our faculty of cognition” in that natural laws are suited “for human understanding in its necessary business of finding the universal for the particular that is offered to it by perception and then further connection in the unity of the principle for all that is different” (5:186). That is, the otherwise “happy accident” that nature is suited to be understood as a systematic whole is required, a priori, as a “purposiveness in relation to the cognitive faculty of the subject” (5:185) that is assumed in every act of the regulative power of judgment. We treat nature as something *to be understood*, something that is *suited* to our cognitive faculties. (We do not have to assume that there literally *is* a designer who gave this purpose *to* nature. Nature is “purposive” – conducive to being understood – without literally having comprehension by humans as its “purpose.”)

For Kant, this assumed purposiveness of nature for our cognition is connected to our faculty of feeling pleasure (or displeasure). The connection is, at first, fairly straightforward: “the attainment of every end is combined with the feeling of pleasure,”⁴⁹ so if reflective judgment gives an a priori aim valid for everyone, “then the feeling of pleasure is also determined through a ground that is a priori and valid for everyone” (5:187). Specifically, since understanding particulars in terms of general laws and “bringing heterogenous laws of nature under higher . . . laws” are demands of reflective judgment made possible through an assumed “purposiveness of nature for our understanding,” “if we succeed in this accord of such laws . . . , pleasure will be felt” (5:187-8). When the paleontologist studying a strange fossil is finally able to classify that fossil as a distinct species falling under some more general genus, she experiences pleasure at this success. The a priori principle of reflective judgment that makes possible the search for systematicity in our understanding of nature thus also provides the first guide to a transcendental anthropology of feeling, since it proposes a necessary *end* for all human beings – unifying particulars under increasingly general laws – the attainment

⁴⁸ Kant does not of course think that there is a rule for the application of a concept. His point is that judgment is a specific power distinct from the power of understanding rules and distinct from the power (reason) of deriving one rule from another. The power of judgment is the power to apply rules to particulars, and this power to apply is not identical to the rule that is applied. But there must still always be some rule that is applied. In the case of determining judgment, this rule is given by the content of the concept.

⁴⁹ Importantly, Kant’s point here is not that one experiences pleasure only when one attains an end (contra Guyer 1979: 81, cf. Allison 2001: 56), but that whenever one attains an end, one experiences pleasure.

of which is a necessary and universal basis of pleasure for human beings. The necessary presumption of strictly contingent purposiveness in nature is the ground of a necessary pleasure in actually discovering such purposiveness.

This pleasure is not aesthetic because it is both interested and conceptual. The pleasure is interested because it is a pleasure in accomplishing a specific goal. The goal is necessary and given by the structure of our cognitive faculties rather than by particular inclinations, but the pleasure is still an *interested* pleasure. The reason that the pleasure is conceptual is a bit more complicated. Insofar as one experiences pleasure in the purposiveness of the world through the systematization of one's experience, one's pleasure is pleasure *in empirical concepts*. And in these cases, the central dilemma of aesthetic judgment does not arise. The scientist who wants to explain why the fossil brings her pleasure can explain that this particular, which did not seem to fit into a systematic whole, can be understood according to such-and-such an empirical concept (Australopithecus, say, or Euleptes gallica). And even if one does not fully share in her pleasure – pleasure of discovery depends at least in part on working for it oneself – one will be able to fully appreciate why pleasure is called for. In that sense, one can explain one's pleasure by reference to (empirical) concepts, and it is not a purely aesthetic pleasure.

In principle, objects in the world might be purposive only in that they possess a general conduciveness to be understood. But in fact, Kant suggests the possibility of a different sort of purposiveness, a purely “subjective” or “aesthetic” purposiveness, “the purposiveness of a thing . . . represented in perception . . . that precedes the cognition of an object, which is immediately connected with it *even without wanting to use the representation of it for a cognition*” (5:189, emphasis added). With the exception of the final clause, this description could apply to any reflective judgment, wherein one must *first* see an object (or a representation of it) as purposive and only *then* subsume it under general concepts. In this case, however, the purposiveness of the object is recognized without either subsuming the object under concept or aiming for such subsumption. “The object is called purposive in this case *only* because its representation is *immediately* connected the feeling of pleasure” (5:189, emphasis added). Such a representation would be “an aesthetic representation of purposiveness” (5:189). So far, however, this extension of the notion of purposiveness seems both arbitrary and mysterious. It is arbitrary because, as Kant immediately notes, it is still an open “question . . . whether there is such a representation of purposiveness at all” (5:189). And it is mysterious because it is not clear what such a representation would be purposive *towards* nor how any representation could be *immediately* connected with pleasure in this way.

At this point, Kant's dilemma of aesthetic judgment has an important role to play. Just as Kant justifies the a priori categories of understanding as conditions of possibility of experience, and the postulate of freedom as a conditions of possibility of moral obligation, so here he uses the dilemma of aesthetic to show the nature and

necessity of aesthetic representations of purposiveness.⁵⁰ Along with his contemporaries, Kant takes for granted that there *are* normative judgments of feeling (“good taste”). Kant’s transcendental anthropology must look for the conditions of possibility of such judgments. They are problematic, he argues, because they must be both subjective and universal. But having described the role of purposiveness in reflecting judgment that aims for systematic, empirical knowledge, Kant presents an account of what an immediately-felt, non-conceptual representation of an object’s purposiveness would have to be. And it turns out that such representations are precisely what would make possible normative claims that are both universal and subjective.

In particular, for Kant, an immediately-felt, non-conceptual representation of an object’s purposiveness would have to be a recognition of the suitability of an object not to any particular concept or concepts, but simply to humans’ cognitive faculties *in general*.

If pleasure is connected with the mere apprehension of the form of an object . . . without relation of this to a concept for determinate cognition, then . . . the pleasure can express nothing but [the form’s] suitability to the cognitive faculties that are in play in the reflecting power of judgment, insofar as they are in play, and thus merely a subjective formal purposiveness of the object. (5:189-90, cf. 5:222, 217)⁵¹

The claim that, in the form of the beautiful, one’s cognitive powers are “in play” is central to Kant’s account. The “play” of these cognitive powers can be contrasted both with the *work* that such powers do (when, for example, reflecting judgment develops empirical concepts or unifies diverse laws under more general ones) and with a possible *conflict* between such powers (such as when one’s perceptions *resist* being brought under general concepts).⁵² For the feeling of beauty, the relevant cognitive powers are the imagination and understanding; when these powers play freely together, one feels aesthetic pleasure. But what precisely is this “free play”? While there is substantial

⁵⁰ See Karl Ameriks, “How to Save Kant’s Deduction of Taste” *Journal of Value Inquiry* 16 (1982): 295-302 (reprinted in *Interpreting Kant’s Critiques*, Oxford: 2003: pp. 285-307), where Ameriks defends Kant’s argument as a regressive one, akin to the argument of the first *Critique*.

⁵¹ The consciousness of the merely formal purposiveness in the play of the cognitive powers of the subject in the case of a representation through which an object is given is the pleasure itself, because it contains a determining ground of the activity of the subject with regard to the animation of its cognitive powers, thus an internal causality (which is purposive) with regard to cognition in general, but without being restricted to a particular cognition, hence it contains the mere form of the subjective purposiveness of a representation in an aesthetic judgment. (5:222)

Now if the determine ground of the judgment on this universal communicability of the representation is to be conceived of merely subjectively, namely without a concept of the object, it can be nothing other than the state of mind that is encountered in the relation of the powers of representation to each other insofar as they relate a given representation to *cognition in general*. (5:217)

⁵² See Allison 2001: 49-50. We will see another important case of conflict between powers when we turn to the sublime, below.

disagreement amongst commentators,⁵³ the general idea can be gleaned from Kant's examples of beautiful objects:

Flowers are free natural beauties. Hardly anyone other than the botanist knows what sort of thing a flower is supposed to be; and even the botanist, who recognizes in it the reproductive organ of the plant, pays no attention to this natural end if he judges the flower by means of taste. Thus the judgment is not grounded on any kind of perfection, any internal purposiveness to which the composition of the manifold is related. Many birds (the parrot, the hummingbird, the bird of paradise) and a host of marine crustaceans are beauties in themselves, which are not attached to a determinate object in accordance with concepts . . . but are free and please for themselves. Thus designs *à la grecque*, foliage for borders or on wallpaper, etc., signify nothing by themselves: they do not represent anything, no object under a determinate concept, and are free beauties. One can also count . . . musical fantasias (without a theme), indeed all music without a text. (5:229).

All of these examples refer to objects that inspire continuous reflection without getting to any determinate knowledge. Unlike cold, clearly conceptualizable forms – such as an equilateral triangle (see 5:241) – that give no room to the imagination to examine them in new ways, a flower stimulates a constant redirection of attention from one aspect of its form to another, a constant attempt to reassemble the visually presented material with different emphases. But in contrast to a merely chaotic mish-mash of stimuli, the diverse perspectives that one can take on a flower are all orderly; the understanding is given constant encouragement to find patterns and generalities in the representations of the object. Moreover, the activities of imagination and understanding do not merely take place side-by-side; they are “reciprocally expeditious” (20:224). Finding patterns in one way of looking at a flower facilitates the re-presentation of the flower in yet another way, which leads to the recognition of a new order, and so on. One can continuously contemplate beautiful flowers, birds,⁵⁴ and musical improvisations, constantly reinterpreting them in the light of new “imaginative” ways of pulling together one’s impressions.

The purposiveness of beautiful objects is thus not a purposiveness towards the ends that our cognitive faculties set for themselves – increasing knowledge of the empirical world – but a purposiveness towards the *activity* of those powers themselves. For Kant, however, pleasure is a feeling of “the agreement of an object with the productive power[s] of the soul” (29:894); the “animation of [the] cognitive powers” of imagination and understanding gives rise to a pleasure, which “is itself” the consciousness of the purposiveness of the beautiful object (5:222). Beautiful objects are *pleasurable*, and because this pleasure lies in the *mere* animation of one’s cognitive powers, and not any *end* brought about by those powers, it is *disinterested*. This animation of cognitive powers is the effect of the mere representation of the object, without – and

⁵³ References xxx and briefly summarize the key issues of debate.

⁵⁴ Crustaceans, too. See <http://www.refkeeping.com/issues/2007-05/as/index.php>, accessed June 4, 2009.

even in contrast with – any determinate cognition of the object, so one’s judgment that the object is beautiful is *non-conceptual* and thus *subjective*. But – and this is Kant’s key move – because the subjective basis of one’s judgment is the free play of cognitive powers *that all human beings share*, one can legitimately expect that *any* human being *should* feel pleasure at the representation of the beautiful object. Because the judgment that an object is beautiful is a judgment that the object is *purposive for one’s cognitive powers*, and because human beings share those cognitive powers in common, an aesthetic judgment carries with it a commitment to universality. (Of course, one might still get aesthetic judgments *wrong*. One’s pleasure in an object might only *seem* to be due to disinterested, non-conceptual contemplation. In that case, one might mistakenly call beautiful what is really agreeable or good.)

The purposiveness that grounds the subjective universality of aesthetic judgments of beauty also provides the basis for truly *free* pleasure. “among all . . . kinds of satisfaction only that of the taste for the beautiful is a disinterested and *free* satisfaction; for no interest, neither that of the senses nor that of reason, extorts approval” (5:210, cf. 5:354). As it typical for Kant’s transcendental anthropology, *freedom* in the context of pleasure is normatively governed. Just as free cognition is governed by a priori categories and forms of intuition, and free volition is governed by a categorical imperative, so the free experience of pleasure is governed by a principle of purposiveness by virtue of which one judges an object to be beautiful and hence worthy of free pleasure when it properly enlivens one’s imagination and understanding. Moreover, the free pleasure in beauty is a particularly *human* sort of pleasure: “Agreeableness is also valid for nonrational animals; beauty is valid only for human beings . . .; the good is valid for every rational being in general” (5:210).

But although aesthetic pleasure is free even from moral considerations, it provides an important “mediating concept” between morality and our sensuous nature (5:196, cf. 5:176). Aesthetic pleasure both provides an “analogy” to morality (5:353) and “prepares us” for it (5:267). With respect to the former, the universality, freedom, disinterested pleasure, and non-conceptual basis of aesthetic pleasure all correspond to similar traits in the case of the moral law. And with respect to the latter, feeling pleasure in beautiful objects proves to human beings that we are not merely the playthings of our instincts and inclinations. Although aesthetic pleasure is not itself motivational (since it is disinterested), this possibility of disinterested pleasure reinforces the categorical demands of morality with a subjective (and pleasing) basis for thinking that one can be sufficiently free from inclination to meet those demands. Moreover, the subjective universality of taste promotes the development of cultures of taste, within which one seeks to conform one’s own aesthetic judgments to those of others. One’s effort to resist what Kant will later call “aesthetic egoism” is implicit in the universal nature of aesthetic judgment itself, and this development into a tasteful, social being prepares the way for overcoming the “moral egoism” that treats one’s own desires as more important than those of others.

In an important sense, then, Kant's account of the feeling of pleasure in beautiful objects completes his transcendental anthropology. With this "critique of aesthetic judgment," Kant has shown that the entirety of human mental life – cognition, volition, and *feeling* – is susceptible to transcendental investigation. Like cognition and volition, human feeling is normative and one can investigate the conditions of possibility of this normative structure from-within. While cognition is governed by a priori principles of the understanding and volition by an a priori principle of reason, feeling is governed by an a priori principle of *judgment*: the principle of purposiveness. Moreover, Kant's account of beauty solves the problem of bridging the gap between nature and freedom in both the cognitive and volitional dimensions. With respect to cognition, the experience of beautiful objects involves reflectively judging about objects in the world and feeling the purposive suitability of this world to our cognitive capacities. This purposiveness of the world for reflective judgment provides a ground (albeit a subjective one) for regulating the investigation of nature in accordance with an assumption of its suitability for systematic understanding, thus bridging the gap between the *Critique of Pure Reason's* assurance that the world would conform to certain general structures of human cognition and the need to be able to expand the scope of knowledge systematically. With respect to volition, the experience of beautiful objects reveals, in the most subjective dimension of human existence, a universality and autonomy that is analogous to and preparatory for moral choice.

Kant *could* have ended his transcendental anthropology with his account of the beautiful, and there is some evidence that he at first *intended* to add little more than this (see xxx), but in fact Kant went on to add two important dimensions to his *Critique of Judgment*: a theory of the sublime, and an account of teleological judgment. The account of the sublime is a natural addition to Kant's critique of aesthetic judgment. Especially with the publication of Edmund Burke's *A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and Beautiful*, the issue of the relationship between aesthetic judgments of beauty and of the sublime was a hot topic in the 18th century. Like the experience of beauty, the experience of sublimity is something with which human beings find ourselves. When contemplating the Milky Way, or St. Peter's in Rome, or a mountain landscape, one can have a distinctive sort of experience of not being able to "take it all in." And while this inability to take it all in involves a certain kind of frustration, can also be strangely satisfying. Similarly, when one experiences "[b]old, overhanging, and, as it were, threatening rocks, thunderclouds towering up in the heavens, . . . volcanoes with their all-destroying violence, . . . [or] the boundless ocean set into a rage," one can feel like one's "capacity to resist [is] an insignificant trifle in comparison with their power" (5:261). Again, however, this painful frustration of one's powers can be conjoined with a strange and quiet sort of pleasure. Kant had "observed" and discussed these "sublime" feelings in detail almost thirty years before writing the *Critique of*

Judgment,⁵⁵ but not until this work does he try to give an account from-within of how such a feeling could be warranted.

In the *Critique of Judgment*, Kant transforms his early, casual account into a detailed theory of the relationship between these two forms of aesthetic judgment. The beautiful, for Kant, is the unambiguously pleasurable feeling of the harmony of our imagination and *understanding*. The sublime, by contrast, is the deeply ambivalent feeling of the *disharmony* between our imagination and our *reason*. There are important similarities between the two:

The beautiful coincides with the sublime in that both please for themselves. And further in that both presuppose neither a judgment of sense nor a logically determining judgment but a judgment of reflection: consequently the satisfaction does not depend on a sensation, like that in the agreeable, nor on a determinate concept, like the satisfaction in the good . . . Hence both sorts of judgments are also *singular*, and yet . . . universally valid in regard to every subject, although they lay claim merely to the feeling of pleasure. (5:244).

The sublime, however, is an odd sort of universal pleasure because “that which . . . excites in us the feeling of the sublime may . . . appear in its form to be *contrapurposive* for our power of judgment, *unsuitable* for our faculty of presentation, and as it were *doing violence* to our imagination, and yet is judged all the more sublime for that” (5:245, emphasis added). Judgments of the sublime involve recognizing *disharmony* between the objects and our powers. In cases of what Kant calls the “mathematically sublime” (5:248f.), where we can’t “take it all in,” the disharmony is between human reason, which seeks “the comprehension of every appearance . . . into the intuition of a whole” (5:257), and the imagination, which is “inadequate” for this task (5:252). In cases of what Kant calls the “dynamically sublime” (5:260f.), the disharmony is between our natural capacities as a whole and a powerful, threatening object or situation. But if the beautiful pleases through the harmony it makes evident amongst our cognitive powers, how can the sublime, which makes evident a *disharmony* amongst those same powers, also please?

With the sublime, an initially painful disharmony gives rise to a feeling of pleasure when one recognizes in oneself a capacity that outstrips nature itself. With mathematically sublime objects, one sees that the demands of one’s (theoretical) reason outstrip nature itself, so that one seeks a unity in nature that cannot be satisfied by anything one’s senses can provide. And with the dynamically sublime, one’s recognition of the frailty of one’s natural life can reveal that one’s natural life does not exhaust who one is, that one’s human nature includes moral demand to have a good will that no natural forces can undermine. In both cases, the experience of certain natural objects gives rise to a feeling of one’s own transcendence over nature itself. Thus, Kant claims, “true sublimity must be sought only in the mind of the one who judges, not in the

⁵⁵ *Observations on the Feeling of the Beautiful and the Sublime* (1762), available in *Observations on the Feeling of the Beautiful and the Sublime and other writings* (ed. Patrick Frierson and Paul Guyer) xxx.

object in nature . . . That is sublime which even to be able to think of demonstrates a faculty of mind that surpasses every measure of the senses” (5:256, 250). In reflecting on certain objects, one comes to recognize a disharmony that is caused by the superiority of one’s humanity over the sensible, natural world. While this disharmony causes an initial sort of pleasure, the *source* of that disharmony – one’s own transcendent reason – causes a quiet and ambivalent, but still intense and pleasurable, feeling of self-esteem.

The feeling of the sublime completes Kant’s account of aesthetic pleasure and supplements his treatment of the beautiful in three important ways. First, insofar as Kant aims to give a complete transcendental anthropology, he must account for all of the ways in which humans’ feelings of pleasure can be governed by a priori norms; since experience of the sublime is genuinely felt within human beings, it must be accounted for. Second, the sublime provides an important balance to the contribution of the beautiful to Kant’s anthropology of cognition, corresponding to the role that the transcendental dialectic plays in Kant’s *Critique of Pure Reason*. There Kant showed that while the understanding structures our experience of the world, reason imposes demands that transcend any possible experience. In the beautiful, we feel the conduciveness of the world to human understanding; in the experience of the sublime, we feel how reason imposes demands that transcend the world. Finally, and perhaps most importantly, the sublime provides a bridge between nature and freedom that is importantly different from that of the beautiful. With the beautiful, humans’ experience of fit between themselves and nature makes us aware of a free, disinterested, universal capacity for pleasure that is analogous to moral demands. With the sublime, especially the dynamically sublime, humans directly feel their moral dignity. The experience of the sublime involves feeling precisely the same sort of respect for oneself that is constitutive of moral motivation. Thus while in one sense, the sublime reflects the disconnect between oneself and nature, it also marks a bridge from an experience of nature that is not itself moral to a respect for oneself that plays a central role in moral motivation.

With his account of the sublime, there is an important sense in which Kant *has* completed his transcendental anthropology of feeling. The rest of the *Critique of Judgment* does not directly address humans’ faculties of feeling at all. But for Kant, while “[i]t is . . . in taste . . . [that] the power of judgment reveals itself as a faculty that has its own special principle and thereby makes a well-founded claim to a place in the general critique of the higher faculties of cognition” (20:244), this a priori principle can be applied more widely than *merely* to govern aesthetic feeling:

once the capacity of the power of judgment to institute a priori principles for itself is granted, then it is also necessary to determine the scope of this capacity, and for this completeness in critique it is required that its aesthetic faculty be recognized as contained in one faculty together with the teleological and as resting on the same principle, for the teleological judgment about things in nature also belongs, just as much as the aesthetic, to the reflecting power of judgment. (20:244)

Already, we have discussed one important non-aesthetic dimension of the power of judgment. Nature's purposiveness to our cognitive powers is not limited *merely* to free play but also enables an increasingly systematic *determinate* understanding of nature. But in the "critique of teleological judgment" (5:357) Kant goes further. Once "teleological judging" is "drawn into our research into nature" (5:360), one can consider the possibility of "objective purposiveness," wherein certain objects in nature are understood as natural purposes. Importantly, nature might have turned out to include objects purposive *for our cognitive powers* but not *objective purposes*.⁵⁶ The laws of physics and chemistry are merely mechanical, even though they show that matter is purposive for our cognitive powers because it can be subsumed under general laws. But Kant argues that we can think of a different sort of purposiveness of nature, what he calls "objective purposiveness":

Experience leads our power of judgment to the concept of an objective . . . purposiveness, i.e., to the concept of an end of nature, only if there is a relation of the cause to the effect to be judged which we can understand as lawful only insofar as we find ourselves capable of subsuming the idea of the effect under the causality of its cause as the underlying condition of possibility of the former. (5:366-7)

For objective purposiveness, the "natural laws" under which we subsume a given phenomenon depend upon thinking of causes of that phenomenon as *for the purpose of* their effect. For example, if one can make sense of "rivers" only by appeal to their role in "carry[ing] with them all sorts of soil helpful for the growth of plants" (5:367), then rivers have an objective purposiveness.⁵⁷ More importantly, if one can make sense of an organized being only as "a thing . . . [that] is cause and effect of itself" (5:370), then this being will be a "natural end." When one understands the motion of the heart in terms of its functional role in promoting the circulation of blood, and the circulation of blood in terms of promoting the life functions of an animal, and these life functions as in turn ensuring the continual motion of the heart, one interprets an animal teleologically, in terms of its purposiveness. When, further, one sees an individual animal as both the effect of its species and the cause of the continuation of the species, one interprets the animal teleologically; it exists *for* the propagation of the species (and vice versa).

In theory, there might not be "natural ends," but in fact one finds self-propagating organized beings in the world "which cannot be explained through [mechanism] alone" (5:374). The result is that human beings are entitled, and even required, to posit a principle for judging organized (biological) beings: "*An organized*

⁵⁶ "Now if nature showed us nothing more than this logical purposiveness, we would indeed already have cause to admire it for this, since we cannot suggest any ground for this in accordance with general laws of the understanding; only hardly anyone other than a transcendental philosopher would be capable of this admiration, and even he would not be able to name any determinate case where this purposiveness proved itself *in concreto*, but would have to think of it only in general." (20:216)

⁵⁷ In this case, of course, the purposiveness is "relative;" the river can be understood as purposive for plants (or for humans who eat the plants) only if plants (or humans) are themselves, in some sense, natural ends. Thus Kant argues that "external purposiveness (advantageousness of one thing for another) . . . justifies no absolute teleological judgments" (5:368-9).

product of nature is that in which everything is an end and reciprocally a means as well.” Or, in less technical lingo, “Nothing is in vain, purposeless, or to be ascribed to a blind mechanism of nature” (5:376). Importantly, for Kant, these principles are purely “regulative,” mere heuristics “for guiding research into objects of this kind” (5:376). But as heuristics, they are

indispensably necessary . . . In fact, [anatomists of plants and animals] could just as little dispense with this teleological principle as they could do without the universal physical principle, since, just as in the case of the abandonment of the latter there would remain no experience at all [as shown in the first *Critique*], so in the case of the abandonment of the former principle there would remain no guideline for the observation of a kind of natural thing that we have conceived of teleologically as an end. (5:376)

For Kant’s transcendental anthropology, the addition of these teleological principles has two important implications. First, it allows a limit to the mechanistic explanation that the first *Critique* justified. While Kant insists that *in principle* everything in nature is explicable mechanistically and even that we are required to explain nature mechanically as much as possible (5:379, 429), he concedes that for humans studying the living world, mechanistic explanations will often not be possible. Second, Kant shows here a willingness to introduce new principles for judgment on the basis of empirical discoveries. The principle of objective purposiveness *precedes* and *guides* empirical research; biologists *assume* purposiveness prior to finding the specific purposes of particular aspects of organized beings. But this assumed purposiveness is itself the result of discovering *through experience* that certain beings in nature can only be understood (by us) in this way.

The need to investigate organized (living) things in accordance with a principle of purposiveness also gives rise to two further implications that will prove important for Kant’s anthropology as a whole. The first of these will be discussed in the next chapter. Briefly, just as Kant’s *Critique of Pure Reason* created a space for an empirical anthropology that views human beings as empirical objects subject to natural laws, his *Critique of Judgment* makes clear that, like other living things, certain aspects of human nature will be irreducible to strict mechanism, requiring instead teleological explanation. The second implication is discussed in detail in the *Critique of Judgment* and constitutes the most important contribution of that work to the third key question of philosophy: “What may I hope?”

After explaining that organized beings in general must be understood as natural ends, Kant explains,

It is therefore only matter insofar as it is organized that necessarily carries with it the concept of itself as a natural end, since its specific form is at the same time a product of nature. However, this concept necessarily leads to the idea of the whole of nature as a system in accordance with the rule of ends . . . (at least in order to test natural appearance by this idea). (5:378-9)

Once one finds the fruitfulness of studying living beings in accordance with the principle that “Nothing is in vain,” it is natural to extend this heuristic principle to studying nature as a whole. This immediately yields fruitfulness in the form of a scientific approach to ecology and interdependence between organisms, as one comes to see, for example, how plants nourish animals that in turn provide “for the human being, for the diverse uses which his understanding teaches him to make of all these creatures,” but also how, in turn, “plant-eating animals exist in order to moderate the excessive growth of the plant kingdom” and humans and other meat-eaters exist to keep the plant-eaters in check (5:426-7). But Kant also insists that this study of nature naturally leads one to think about what could be the “final end” of nature as whole. “A final end,” for Kant, “is that end which needs no other as the [teleological] condition of its possibility” (5:434), that is, something that we can see as being a self-sufficient *end-in-itself*. For organisms in general, one can understand the parts of the organisms as teleological ordered towards the whole, and even the individual organism as ordered towards the species, but one can still ask “why do these creatures exist?” (5:426). Answering this ecologically only pushes the question back further, “but why do *those* creatures exist?” Naturally, Kant argues, we want an answer that justifies the whole world.

Given Kant’s transcendental anthropology of volition, the answer is both obvious and problematic. The answer is obvious in that Kant has already shown that there *is* something that is an end-in-itself: *humanity*. But the humanity that is an end-in-itself requires the transcendental freedom that grounds the possibility of a good will. And *that* cannot be an end *of nature* because neither transcendental freedom nor the good will are objects in nature; Kant establishes freedom and morality from a practical rather than empirical standpoint, as “things-in-themselves” rather than “appearances.” Thus Kant insists that despite seeking a final end of nature, “if we go through the whole of nature, we do not find in it, as nature, any being that can claim the privilege of being the final end of creation” (5:426).

Fortunately, however, once we know that human beings as transcendently free choosers are the *final* end of nature, we can look for an “*ultimate*” end of nature that would identify “that which nature is capable of doing in order to prepare him for what he must himself do in order to be a final end” (5:430). Identifying this “ultimate end of nature” provides the basis for a rational hope that nature will cooperate with our moral vocation. In the *Critique of Judgment*, Kant focuses on the regulative principles that can guide our empirical study of nature. Given that humans’ *final* end is moral, Kant suggests that the *ultimate* end of nature, as that which nature can do to support our moral freedom, must be the human being’s “aptitude for setting himself ends at all and . . . using nature as a means appropriate to the maxims of his free ends in general” (5:431). Thus nature as a whole tends towards human beings cultivation in “skill” at using things for one’s purposes, the “discipline” whereby we rise above our inclinations, civil societies that establish objectively right relationships amongst people, and even the emergence of cultures of taste, within which “beautiful arts and sciences” flourish

(5:432-3).⁵⁸ The *details* of Kant's account of human beings as the ultimate end of nature, including the empirical evidence that emerge from (and in turn support) his regulative principles, emerge elsewhere and will be discussed in chapter four. But Kant's *Critique of Judgment* shows how purposiveness as the principle of regulative judgment not only grounds aesthetic judgments but even leads, through its application to biology and ecology, to a conception of human beings as ultimate ends of a purposively ordered nature.

In the end, Kant's *Critique of Judgment* provides a complete transcendental anthropology of the faculty of feeling and the power of judgment that provides that faculty with its regulative principle. As an analysis of feeling from-within, the *Critique* shows how there can be non-conceptual normative standards for judgments of taste, and it reveals an analogy to morality in the most sensuous aspect of human nature: our feelings of pleasure and pain. In its further analysis of purposiveness in the study of nature, this *Critique* not only justifies the assumption of order in nature but even shows how the assumption of purposiveness plays a necessary role in regulating humans' study of objects with a certain kind of complexity. Throughout this *Critique*, Kant also emphasizes that purposiveness as a principle of judgment is merely *regulative*, and in that sense it is more intimately human than either the cognitive principles of the first *Critique*, which constitute the structure of the empirical world, or the moral principles of the second, which apply to all rational agents. In the end, though, these intimately human principles provide a foundation for answering the question "What may I hope?" both affectively and rationally. In aesthetic pleasure, we legitimately feel hopeful in our cognitive strivings for systematic understanding of the world and in our moral aspirations for disinterested, universally-justifiable choices. And in our understanding of nature as a teleologically-ordered whole, we look for (and find⁵⁹) evidence that nature as a whole cooperates with our highest moral vocation.

IV. Conclusion

In the end, Kant's three *Critiques* present a picture of human beings as finite but free knowers, actors, and feelers. Human knowledge is constituted by passively received intuitions that are conceptualized by an understanding that spontaneously (that is, freely) imposes categories to cognize objects. Human action involves subordinating subjective and therefore finite maxims to an autonomous moral principle. And aesthetic pleasure arises from the free play of faculties that testify to our finitude. Moreover, the transcendental anthropology of volition in particular provides a (practical) proof that one is a *transcendentally* free "homo noumenon," capable of acting on grounds that are undetermined by empirical causes. The transcendental anthropology of cognition

⁵⁸ Kant emphasizes that human progress, and *not* human *happiness*, is the end of nature (5:431).

⁵⁹ In chapter four, I discuss in more detail the empirical investigations of human history in which Kant finds evidence of such progress.

ensures that the empirical *expression* of one's transcendently free choices will always be a "homo phenomenon," susceptible to empirical description in terms of natural laws (6:417-8). And the transcendental anthropology of feeling shows how the empirically given world supplies material that provokes pleasurable aesthetic feelings that, in different ways, reveal our freedom to us.

Insofar as humans are homo phenomenon, they must be understood in terms of the categories of the understanding and forms of intuition. But even as homo phenomenon, humans are still distinct from merely physical nature in that we are teleologically-ordered biological organisms with particular features, many of which can have important implications for applying the moral law in practical life. Insofar as human beings are free homo noumena, we are both negatively free, in that our (noumenal) choices are not determined by any particular empirical causes, and positively free, in that we are subject to the moral law as the law of our own will (autonomous). By virtue of our freedom, we are worthy of respect and hence the proper "end" of moral choice, and we are worthy of "awe" and hence proper objects of sublime feeling.

Most fundamentally, human's noumenal freedom manifests itself in the freedom of choice that underlies moral responsibility, but all three of Kant's critiques deal with human beings as free agents in the sense that humans are subject to normativity not only in action but also in thinking and feeling. All three realms of human life involve laws, norms, and requirements that are not *causal* but nonetheless perceived as binding from-within. Kant's critical works thus not only set up the general framework of phenomenal-noumenal humanity but also specifically address the *noncausal* laws that govern human beings. Unlike the observation-based anthropology that Kant develops elsewhere (and that we'll discuss in the next chapter), this transcendental anthropology is thoroughly normative throughout, a normative account of the human being from-within: an epistemology, an ethics, and an aesthetics.

As it stands, however, this "transcendental anthropology" is incomplete as an overall answer to the question "What is the human being?" The *Critique of Pure Reason* shows that empirical knowledge is possible and that reason pursues more and more completeness of knowledge of the empirical world. Because human beings appear in the empirical world, transcendental anthropology must be supplemented with an empirical anthropology that describes what humans look like "from-without," and since we are both physical and biological, this empirical anthropology should involve both mechanical and teleological explanations. Moreover, Kant's *a priori* moral philosophy requires supplementation by an "empirical part" that will involve "judgment sharpened by experience" to know how the moral law should be applied and how "to provide [it] with access to the human will" (4: 388-9). And finally, while the practical postulates of God and immortality and the general teleology revealed through natural beauty and human biology give *some* basis for moral hope, Kant suggests that we need more detailed *empirical* evidence to support our hope that we can achieve our moral ends (refxxx). It should thus come as no surprise that while Kant was developing his *transcendental* anthropology, he was also engaged in detailed *empirical* studies of human

beings. Such empirical study is necessary to complete his answer to his question “What is the human being?” and thereby to fully answer his remaining questions: “What can I know [including empirically about human beings]?” “What ought I do [to human beings with the empirical features that we have]?” and “What may I hope [based on the progress human being have made historically so far]?” It is to this *empirical* anthropology, then, that we now turn.

Summary

Further Reading

- Henry Allison, *Kant's Transcendental Idealism*. Classic work on the *Critique of Pure Reason* by top Kant scholar that defined the field for a generation and helped make the two-standpoint reading of Kant's transcendental idealism mainstream.
- ———, *Kant's Theory of Freedom*. Classic work on Kant's moral philosophy and theory of freedom by leading Kant scholar.
- ———, *Kant's Theory of Taste*. Recent work on Kant's Critique of Aesthetic Judgment by leading Kant scholar.
- Paul Guyer, *Kant*. A good but slanted introduction to Kant's Critical philosophy as a whole by a leading Kant scholar. Generally more critical of Kant than either Allison or Wood.
- Paul Guyer, ed. *Cambridge Companion to Kant*. Excellent collection of essays by leading Kant scholars on various aspects of Kant's transcendental philosophy.
- Christine Korsgaard, *Creating the Kingdom of Ends*. Very good though somewhat revisionary introduction to Kant's moral philosophy by the top neoKantian moral theorist alive today.
- Michelle Grier, *Kant's Questions: What Can I Know?* (Routledge xxx). Good introduction to Kant's *Critique of Pure Reason*, focusing on the question, “What can I know?” and placing Kant's answer in a historical and contemporary context.
- Julien Wuerth, *Kant's Questions: What Ought I Do?* Good introduction to Kant's moral philosophy, focusing on the question, “What can I know?” and placing Kant's answer in a historical and contemporary context. Takes on Korsgaard's reading of Kant in several respects.
- Allen Wood, *Kant* (Blackwell: xxx). Very good short introduction to Kant's philosophy as a whole by a leading Kant scholar. Wood is also attentive to the empirical dimensions of Kant's philosophy.
- Rachel Zuckert, *Kant on Beauty and Biology*. Very good treatment of the *Critique of Judgment*, one of the only books to discuss the work as a coherent whole.