

INTRODUCTION

[W]hat is man's ultimate nature? We keep returning to the subject with a sense of hesitancy and even dread. For if ... the mind can somehow be explained as the summed activity of a finite number of chemical and electrical reactions, boundaries limit the human prospect—we are biological and our souls cannot fly free. If humankind evolved by Darwinian natural selection, genetic chance and environmental necessity, not God, made the species ... However much we embellish that stark conclusion with metaphor and imagery, it remains the philosophical legacy of the last century of scientific research. No way appears around this admittedly unappealing proposition. It is the essential first hypothesis for any serious consideration of the human condition. (Wilson 1978/2004: 1-2)

If we want to discover what the human being amounts to, we can only find it in what human beings are: and what human beings are, above all other things, is various. It is in understanding that variousness – its range, its nature, its basis, and its implications – that we shall come to construct a concept of human nature ... To be human here is thus not to be Everyman; it is to be a particular kind of human being, and of course human beings differ ... [I]t is in a systematic review and analysis of [different ways of being human] – of the Plain's Indian's bravura, the Hindu's obsessiveness, the Frenchman's rationalism, the Berber's anarchism, the American's optimism – that we shall find out what it is, or can be, to be a [hu]man. (Geertz 1973: 52-3)¹

[T]here is at least one being [the human being] whose existence comes before its essence, a being which exists before it can be defined by any conception of it. ... What do we mean by saying that existence precedes essence? We mean that man first of all exists, encounters himself, surges up in the world – and defines himself afterwards. If the human being ... is not definable, it is because to begin with he is nothing. He will not be anything until later, and then he will be what he makes of himself. Thus, there is no human nature The human being simply is. Not that he is simply what he conceives himself to be, but he is what he wills, and as he conceives himself after already existing – as he wills to be after that leap towards existence. The human being is nothing else but what he makes of himself. (Sartre 1993: 15)

What is the human being? The three quotations with which this introduction begins lay out three alternatives: E.O. Wilson, one of the preeminent sociobiologists of the 20th century, sees

¹ Throughout this quotation and the following one, I've changed "man" and "men" here to "the human being" or "human beings."

the parameters for the answer given by biology. Humans are animals with a particular structure that has evolved over millions of years. We are biological beings, and what we need in order to better answer the question “What is the human being?” is better biology, a more detailed description of how we, as humans, are like and unlike other animals that inhabit the earth. Clifford Geertz, five years earlier, articulated a different conception of human beings, one at the core of the “human” science of anthropology. For Geertz, there is no answer to the question “What is *the* human being?” because we are not preeminently *biological* organisms, but *cultural* ones; since there is not one human culture, there is not one kind of human being. What we need is not better biological description, but more widely-ranging, deeply-investigating studies of human variety. Rather than looking for a theory of human nature, we should seek a catalog of human ways of life. Alternatively, perhaps the problem of finding human nature is more difficult; perhaps, as the existentialist philosopher and literary author Jean-Paul Sartre argues, human beings are “condemned to be free” (Sartre 1956:568). Rather than trying to *discover* what human beings are, we should *make* human nature by free choices. Rather than looking as scientists or anthropologists at what human beings happen to be, we should take the role of architects of possibility, whether as artists (literary or otherwise) imagining and thereby creating new human possibilities, as political or social activists changing the social landscape, or simply as acting individuals creating human nature through our daily choices.

The world we live in today is one within which these approaches to the question “what is the human being?” cannot be ignored. Scientific knowledge about our biological nature – from the coding of the human genome to the mapping of brain activity – has made it clearer than ever that humans operate with biological constraints. As Wilson rightly points out, our knowledge of how our biology is like and unlike that of other animals cannot be ignored in any serious consideration of human nature. At the same time, as the world becomes increasingly interconnected, human diversity, even if actually diminishing, is becoming more apparent and more relevant to more and more people across the globe. Protestant Christians in the United States cannot afford to be wholly ignorant of the cultures of rural Muslims in Afghanistan or atheists in China or Buddhists in Sri Lanka. Throughout the world, diversity is literally on one’s doorstep, as Catholic Filipinos work in Korea and Dubai, Muslim immigrants serve in European Parliaments, and Chinese businessmen set up shop in Africa. The awareness of this diversity requires dealing with the fact that human nature is diverse. Finally, the increased power over our world and ourselves that comes from scientific, technological, and economic progress along with the awareness of the range of human possibilities that comes from seeing other cultures gives rise to an ever-more-acute sense that human nature really is up to us, that we can make ourselves into whatever we want to be. The current situation requires thinking carefully about what it means to be human.

But what, precisely, is one seeking when one asks “What is the human being?” What is the question that Wilson, Geertz, Sartre, and so many others are trying to answer? Strikingly, *none* of these thinkers – not even the biologist Wilson -- treats the question “What is the human being?” quite like the question “what is oxygen?” or “what is a giraffe?” All of them see the question as one about our *prospects*, as one not merely about the structure of our brain or society, but about the *implications* of that structure for human choices, for what we should do with ourselves. All recognize that the question “what *is* the human being?” is also, and fundamentally, about what is *important* about us.

When we understand it in this way, we can see why this question was central for Kant, why Kant would insist, “[t]he greatest concern of the human being is ... to rightly understand what one must do in order to be a human being” (20:41). Knowing what it is to be human is – for Wilson, Geertz, and Sartre no less than for Kant – something worthy of *concern*. Thus Geertz does not simply assert that humans are different, but adds that their differences are *more important* than their similarities, more essential to what it means to be human. All of them recognize a point made by Martin Heidegger, that “The being whose analysis our task is, is always we ourselves. The being of this being is always *mine*” (Heidegger 1953:39). They recognize that asking what a human being is really amounts to asking “who am I?”, “what is most important about me?”, “what do I value about myself?” and even “what do I aspire to be?”

The emphasis on values and aspirations, however, should not blind us to the fact that claims about human prospects and aspirations include descriptions of human beings. Even Sartre, who insists that what human beings *are* can only be answered *after* we make ourselves what we are to be, nonetheless recognizes that we are “condemned to be free”, that freedom is a “human condition” from which we cannot escape. Descriptions of the human condition provide the backdrop for claims about how to act in response to them. Kant, too, recognizes the importance of accurate descriptions of human beings. In part, this is for practical reasons: “The question is which condition suits the human being, an inhabitant of the planet that orbits the sun at a distance of 200 diameters of the sun. Just as little as I can ascend from here to the planet Jupiter, so little do I demand to have qualities that are proper only to that planet ... I do not at all have the ambition of wanting to be a seraph; my pride is only this, that I am a human being” (20:47). One needs to know what human beings *are* to know what we should aspire to be. And for Kant – as, at least, for Wilson and Geertz – human beings are also just very interesting to study. Not only must any practical account of human beings reflect an accurate description of them, but such descriptions are, in their own right, worth pursuing.

At its core, the question “What is the human being?” combines careful description of human characteristics with a normative, aspirational account of what about “us” is or would be truly valuable, an account rooted in the sense that each human questioner has of herself. Answering the question, however, involves clarifying what precisely constitutes a legitimate sort of “description” and also what structure and importance to ascribe to the normative perspective on oneself. And ultimately, as we see in the brief references to Wilson, Geertz, and Sartre, the answer to the question will combine – either implicitly or explicitly – these two aspects.

The main purpose of this book is to lay out Kant’s answer to his question and to situate this answer in the context of contemporary debates about human nature and historical influences that brought us to where we are today. The first part of the book thus focuses on Kant’s answer. An interlude lays out later trends that took the question in different directions. And the final part brings Kant into dialogue with the most important contemporary approaches to human nature, including those of Wilson, Geertz, and Sartre. This short book cannot neither fully detail Kant’s answer nor survey all relevant contemporary approaches. But it introduces key Kantian and contemporary ideas, and “further reading” sections concluding each chapter suggest more detailed treatments of each topic.

1. Kant’s “Anthropologies”

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 (*logos*) of human beings (*anthropos*). Thus Kantian anthropology includes comparisons between different people at different times, but it also includes – and emphasizes – general features of human beings as such. Anthropology is the discipline that answers the question, “What is the Human Being?” That is how the term will be used throughout this book.

But the claim that all of philosophy can be reckoned as anthropology may seem strange for other reasons, as well. While human nature may be a *part* of philosophy, philosophy often deals with questions such as the existence of God or the basic nature of reality that seem to go beyond anthropology, and other disciplines deal (arguably better than philosophy) with important aspects of human nature. In equating philosophy and anthropology, Kant explicitly claims that every really important question that humans can ask, whether about God or substance or basic laws of physics or morals or aesthetics, is fundamentally a question *about* human beings, about what we can know, or should do, or may hope.

A final reason that Kant’s claim to reduce all philosophy to anthropology might seem strange, especially for those accustomed to think of anthropology as an empirical discipline, is that this sort of enquiry seems inadequate to establish the normative claims embodied in the questions of what one *can* (legitimately) believe, or *should* do. Those familiar with Kant’s work may be even more puzzled. At the end of his life, Kant did 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. This *Anthropology* is striking for being deeply empirical, while Kant’s answers to the questions of knowledge, obligation, and hope emphasize that these questions must be answered non-empirically. 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 to 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 and thus cannot articulate the most important dimensions of Kant’s answers to the three questions that, supposedly, can all be “reckon[ed] ... as anthropology.”

Kant articulates different “anthropologies,” different kinds of answer to the question “what is the human being?” Most importantly, he distinguishes between three ways in which one can ask the question and three dimensions of human life to which each of these three ways apply. The dimensions of human life arise from Kant’s description of human mental states as being essentially of three kinds: cognitions (of truth), feelings (of pleasure), and volitions (for various goods). Kant does not ascribe consistent names to his three ways of inquiring, but in this book, I refer to them as “transcendental,” “empirical,” and “pragmatic.” Put very briefly, transcendental anthropology provides normative, from-within accounts of what it’s like to be human, accounts that define how one *should* think, feel, and choose based on what we take ourselves to be doing

when we engage in thinking, feeling, or choosing. Empirical anthropology provides scientific (in a loose sense), observation-based descriptions and categorizations of how observable humans think, feel, and act. And pragmatic anthropology puts these two approaches together, drawing on empirical descriptions to provide advice about how best to satisfy the norms elucidated within transcendental anthropology. Part One of this book unpacks these different Kantian “anthropologies.”

2. What is the human being today?

Kant’s approach to the human being cannot satisfy “the greatest concern of the human being” (20:41) nor achieve the great goal that Kant assigned for it – “to establish the rights of humanity” (20:44) – unless it can be brought into conversation with the dominant approaches to thinking about human beings today. After an Interlude in which I examine the accounts of human beings of five of the most important thinkers of the 19th and early 20th centuries, Part Two cultivates a series of interactions between Kant and the leading contemporary approaches to the question “What is the human being?”

Chapter Seven looks at scientific naturalists such as Wilson, who advocate that the question is best answered by biological or psychological studies of human beings. There are a wide range of such naturalist approaches, so this chapter gives only a relatively small sample of the ways in which philosophers and scientists have sought to use biological or psychological descriptions of human nature – what Kant would call “empirical anthropology” – to *fully* answer the question “What is the human being?” Chapter Eight looks at approaches to human beings that emphasize human diversity, whether in the context of historical changes that make the human being of today different from the human beings of other times or in the context of contemporaneous cultural differences that make human beings in one culture different from those in another. Both of these approaches represent attempts to make what for Kant is only empirical anthropology (or even a subset of empirical anthropology) into the whole, and both approaches not only raise serious problems for Kant but also – as I hope to show – suffer from serious weaknesses that Kant’s anthropology can highlight and alleviate.

Among the most important weaknesses of historicism and naturalism is their failure to take sufficiently seriously what I will call the from-within perspective of transcendental anthropology, and Chapter Nine looks at a philosophical approach to human beings that seeks to take this quite seriously, but with a different result than Kant: existentialism. Existentialists are arguably the most direct *heirs* of Kant’s work in transcendental anthropology, taking a core insight of Kant’s – that human beings are fundamentally free but finite beings – and radicalizing this insight in such a way that the normative weight Kant ascribes to principles of reasoning and action becomes subordinated to – rather than constitutive of – human freedom.

In the end, I argue that while existentialism can reinvigorate and even enrich certain Kantian emphases, it fails to speak to human beings because it fails to provide the right sort of normativity. Chapter Ten then takes up a small sample of contemporary approaches to normativity, beginning with some that are far removed from Kant and ending with two of the most prominent contemporary neo-Kantian philosophers writing today: Jürgen Habermas and Christine Korsgaard. These philosophers provide models for how Kantians today can integrate

and respond to the insights of naturalism, historicism, and existentialism while still developing authentically Kantian conceptions of the human being.

PART ONE: KANT ON THE HUMAN BEING

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 ... 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. “Transcendental Anthropology”

In the introduction, I claimed that Kant’s answer to the question “What is the human being?” has at least three different components. Of these, I will refer to the one that made Kant famous and that he identified with “the field of philosophy” (9:25) as “transcendental anthropology.” The term “transcendental anthropology” is taken from Kant’s handwritten notes, in which he refers to an “anthropologia transcendentalis,” a “self-knowledge of the understanding and reason” that would critique all other sciences, including not only “geometry” and “knowledge of nature” but even “literature ... theology, law” and “knowledge of morality” (RA 903, 15:395). But my use of the term arises from Kant’s insistence that all of philosophy is reducible to “anthropology” (9:25) and his description of each aspect of his philosophy as “transcendental” (see A13/B27; 4:390; 5:113, 266, 270; 6:272; and 8:381). Admittedly, Kant often uses the term “anthropology” for his *pragmatic* anthropology, and he often reserves the term “transcendental” for investigations of the conditions of possibility of *experience* (the topic of only one of his *Critiques*). But he does use both terms in broader ways throughout his works, and “transcendental anthropology” provides a useful term to contrast Kant’s approach to the human being in his a priori philosophical works with the empirical and pragmatic approaches elsewhere. Throughout his philosophical works, Kant answers central philosophical questions in ways that are “anthropological,” but in a distinctive sense of anthropology that I call “transcendental.”

While this transcendental investigation is contrasted, for Kant, with empirical study of human beings, one must be careful not to confuse “transcendental” with “transcendent” and thereby take transcendental anthropology (or philosophy) to refer to some 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, in transcendental anthropology, one knows oneself from-within rather than looking at one’s psychology from the stance of an observer. Transcendental anthropology is a most *immanent* self-knowledge, and hence sharply contrasted with both empirical sciences *and* divine-like transcendence.

The notion of transcendental anthropology as “from-within” is often described in terms of a difference between “first person” and “third person” perspectives, the perspectives of the thinking, feeling, or choosing *subject* and perspectives *on* someone as an object. This way of describing the distinction can be helpful if one avoids thinking of “introspective” states as first person, since “from-within” 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. One observing humans might say that what a person chooses in a particular case is determined by accidental environmental features 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 anger and exhaustion. The next chapter shows how Kant’s *empirical* anthropology focuses on these sorts of causal explanations of behavior. But when one actually choosing, one doesn’t consider these accidental and unconscious influences as bases for choice. One looks for various reasons for action, and even if these reasons 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. Kant’s “transcendental” characterizes the processes of thinking, judging, choice, and aesthetic appreciation from-within.

The from-within perspective 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.” The question whether, say, anger is a “good” *cause* seems misguided; it either is the cause or it is not. But when thinking about behavior (or judgments, or choices) transcendently, one looks at reasons for behavior, and 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 this is the sort of question 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 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 empirical cognition (what we can know). As an experiencer of the world, one can think about what must be the case 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 moral law we find valid within 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 basic mental faculties “*from-within*” that specifically attends to the *conditions of possibility* of *normative* constraints on human beings. The rest of this chapter takes up some details of this transcendental anthropology as it plays out in Kant’s three famous *Critiques of Pure Reason, Practical Reason, and Judgment*. Before turning to those

details, it is worth saying a bit more about the specifics of Kant's conception of the human being in order to see how the *Critiques* hang together as a whole "transcendental anthropology" and thus how "we could reckon all of [philosophy] as anthropology" (9:25). For Kant, human mental states are divided into cognitions, volitions, and feelings. Each 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 in terms of these different human faculties (5:198).²

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	<i>Critique of Judgment</i> (1790)	What may I hope?
Desire/ Volition	Reason	Final End	Freedom	<i>Critique of Practical Reason</i> (1788)	What ought I to do?

II. What can I know?

The Critique of Pure Reason as transcendental anthropology of cognition

Kant's most famous and important work, the *Critique of Pure Reason*, 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 here 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. Kant starts with an interest in the status of traditional metaphysics, which

² 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. As with most of Kant's tidy charts, this one hides many complications (for instance, Kant typically identifies his question "what may I hope?" with his writings on religion and not directly with the *Critique of Judgment*), but it is helpful for a general overview.

involves claims that are “*a priori*” in that they are necessary and thus not based merely on empirical generalizations, but also “synthetic” because they put together concepts to make substantive assertions about the world. But this metaphysics raises “the general problem” of the *Critique of Pure Reason*: “How are synthetic judgments possible a priori?” (B19, cf. *Prol.* 4:276). In answering this question, Kant aims to answer the question “What can I know?” as it applies to the “objective validity” of “a priori concepts” (A xvi), that is, “what and how much can the understanding and reason cognize free of all experience?” (A xvii). Through this transcendental anthropology of cognition, Kant defends a metaphysics that consists in a priori claims about the nature of the world and lays out an epistemology that limits the scope of such claims.

Kant’s answer to the question of the possibility of synthetic a priori knowledge depends upon conceiving of metaphysics as a subset of transcendental anthropology. From the beginning of his *Critique*, Kant makes his radically human-centered metaphysics clear:

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 ... 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 objects must conform to our cognition. (Bxvi)

To move human cognition *into* the center of metaphysics, Kant 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. Previous philosophers – especially during the 17th and 18th centuries – either sought philosophical systems based upon reason alone (rationalism) or sought the ultimate foundations of knowledge in experience (empiricism). 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). 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. There is better hope of showing how a priori synthetic judgments are possible if one assumes instead that the world of objects must conform to the structure of human cognition (B xvi). If the world conforms to our cognition, we can know about the world based on the structure of our cognition rather than by induction from experience.

Kant’s next move 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 humans’ thoughts about objects have two components, an active component by which we *think* about objects, and a passive component by which thoughts are *about objects*: “Thoughts without content are empty, intuitions without concepts are blind” (A50-1/B74-5). Knowledge of an objective world involves receiving “intuitions” from the world and processing them using one’s concepts.

Kant’s appeal to “intuitions” – a technical term for that which is given by 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, humans have knowledge we take *from* the world, such as that there are mountains in the Pacific Northwest of North America, that water freezes, 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. 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 to 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 that goes beyond mere conceptual analysis. Even human receptivity has an a priori structure to which the world must conform, and so there can be 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. In particular, Kant argues that space and time are a priori intuitions that structure all humans' empirical intuitions. We can neither think of the world as non-spatial or non-temporal nor think of external objects without an already-given spatial (and temporal) structure (A22-5/B37-40, A30-2/B46-8). And given space and time as a priori intuitions, Kant can explain the success of geometry (based on space) and arithmetic (time), both of which give synthetic a priori knowledge (A25/B40-1).

Human understanding, like sensibility, has an a priori structure, and after a lengthy, detailed, and controversial defense of a set of a priori "categories" of thought,³ 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 best known) is sufficient to give a sense for his general strategy. Kant defends the principle of cause and effect as one by which human beings structure the objective world: "All alterations occur in accordance with the law of the connection of cause and effect" (B232). Humans experience a changing world, so Kant's argument considers what is necessary in order for a set of perceptions to be considered perceptions of alteration (or, more generally, of something happening). Kant distinguishes merely subjective perceptions from objective experience. To have objective experience, one 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 another category of connection (say, object—property or part—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 thing and one supposes all the parts of a thing to exist at the same time. As an example of a purely subjective sequence, Kant describes the perception of a house, starting with the chimney, then the roof, then the windows, then the door. Here one doesn't suppose that objectively speaking there really is first a roof, then windows, then a door, etc. By contrast, Kant gives the example of a boat, where one perceives a boat upstream, a boat midstream, and a boat downstream. Here one

³ For discussion, see Grier forthcoming, Allison 1983, Ameriks 2003, Guyer 1987, and Longuenesse 1997.

supposes not that these are different parts of a complicated stream-wide boat, but that in reality – that is, objectively – the boat is moving. 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. But given Kant's anthropological orientation, to say that one must order perceptions in a certain way is just to say that the objects of those perceptions must in 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 structure a 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 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 cognitions. And because some of this cognition is of objective succession, we must apply categories of cause and effect to structure the world we experience.

These a priori claims about the world are “only from the human standpoint” (A26/B 42), which can be contrasted with, say, the standpoints of animals and God. Animals – according to Kant – lack “inner sense” and thereby the self-consciousness needed for reflection (see LM 28:276). As a result, animals have “mere sensations” or “intuitions” and thereby no true “experience,” which comes only with the addition of concepts (LL 9:236, 702). Animal “cognition,” if we can call it that, is a mere “analogue of reason,” something purely “immediate” that “cannot be described” (LL 9:236). While animal *intuition* may be like our own (see LM 28:297, 888), animals lack the faculties of understanding and reason that give us the spontaneity to organize our representations into experience. And in the absence of this spontaneity of understanding, even animals' intuitive awareness of space and time will be markedly different from our own (not involving cognition of causal relations, for example). At the other extreme, God has a purely “intellectual intuition” with no passive sensibility at all. God's knowing is all spontaneity without finitude. From God's cognition, “one is careful to remove the conditions of time and space,” and Kant claims that God's knowing will actually be a pure but creative intuition, “one through which the existence of the object is itself given” (B72). Whereas humans' forms of intuition provide a structure within which objects are given *to us*, God's intellectual intuition actually brings objects into existence. And God need not then “think” about such objects, since God fully knows them through the intuition by which they were created.

Unlike mere animals or God, human knowers are free but also finite, capable of cognizing objects given through spatial and temporal intuitions in terms of categories of the understanding applied to those objects. There could, of course, be other free but finite knowers, who would also need to understand an intuitively given world in terms of a priori categories. The

categories, as basic structures of thought itself, constrain any discursive understanding of the world. But Kant seems open to the possibility that other finite rational beings could either share our forms of intuition or have different ones, saying only that “we cannot decide this” (B72).

In the end, Kant’s *Critique of Pure Reason* provides a transcendental analysis of human faculties of sensibility and understanding that elucidates their a priori structure and the contributions of this structure to experience of an objective world. Human beings for whom experience and a priori synthetic judgments are possible are finite beings dependent upon sensibility and also spontaneous free thinkers. And this transcendental anthropology provides both an epistemology that delimits what we can know and a metaphysics of a world that must conform to human cognition. Metaphysics and epistemology turn out, in Kant’s hands, to be reckoned as (transcendental) anthropology.

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 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” (4:469-70). 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 and these necessary ways of cognizing would be a priori synthetic principles of material bodies themselves. What this implies, for Kant, is that even the basic principles of physics are part of transcendental anthropology. Kant even argues that given the existence of matter in motion, one can derive a priori such claims as the conservation of matter (4:541) and Newton’s laws (4:543, 4:554). For Kant, not only the most basic metaphysical claims about the universe, but even Newtonian physics, is transcendental anthropology.

The second further element comes in the second part of the *Critique of Pure Reason*. There Kant turns from human sensibility and understanding to human “reason,” which presents ideals that regulate humans’ pursuit of knowledge by constantly seeking the “unconditioned,” that is, an answer that does not itself require a further explanation. As reason drives humans to learn more and more about their world in a search for the unconditioned, it generates illusions that an unconditioned is there to be found. The second half of Kant’s *Critique of Pure Reason* focuses on the dangers of these illusions, showing how 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. The details of these arguments are unnecessary in this brief account of Kant’s transcendental anthropology, but one discussion is particularly important for Kant’s conception of the human being. Reason, in explaining causality, finds itself seeking a “free” cause that does not itself have a prior cause. Consistent with his general approach to the illusions of reason, Kant warns against assuming that any such free cause could exist in the world, but Kant’s discussion

of freedom also highlights 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 anthropology can make towards a robust metaphysics of nature. 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 cannot provide any theoretical justification for any claims about unexperienceable things. 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. He describes his position as "empirical realism" because its claims (e.g. about causality) are necessarily true of the empirical world but also as "transcendental idealism" because such claims are limited to the empirical world and say nothing about what "things-in-themselves" – apart from human sensibility – might be like.

In the case of freedom, this transcendental idealism does significant work. Consistent with his insights regarding causation, Kant insists that any objective alteration must be the result of causes in accordance with natural laws. But he then asks "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:

[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 ... an 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 causation] and is not itself [an empirical object]. (A539/B567)

The distinction between empirical and intelligible character makes it possible for Kant to defend the possibility of what he 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 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 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 is 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 conditions and therefore unfree, or “libertarians” who believe that (some) events in the world are determined by choices that are not determined by prior conditions. Kant’s position has aptly been called a “compatibility of compatibilism and [libertarian] incompatibilism” (Wood 1984: 74). 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 distinctive 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.

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 ourselves as free. Insofar as we take a scientific-observer standpoint, we must see everything (including ourselves) as causally determined.

On either interpretation, the theory of freedom that Kant lays out in the first *Critique* is presented only as an option that metaphysics can neither establish nor rule out, an “extension” that “even if ... empty, ... we ... can fill through practical data of reason” (Bxxi, cf. A558/B586). And this sort of modesty shows an important positive aspect of Kant’s limitations of metaphysics. After noting how his account of cognition precludes metaphysical proofs about things like God, humans’ immortal souls, and freedom, Kant adds,

this critique is also in fact of positive and very important 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). Kant used 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 this non-cognitive access

in another part of his transcendental anthropology, the transcendental analysis of volition wherein *morality* provides a non-cognitive role for reason in governing human life.

III. What ought I to 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 to do?” While Kant’s transcendental analysis of cognition focused on human beings as free but finite knowers, Kant here thinks about human beings as free but finite agents. In some respects, Kant’s foundational work in moral philosophy – *Groundwork of a Metaphysics of Morals* – might seem specifically to avoid developing ethics as a subset of anthropology: “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, see too 4:410-12, 425). Kant is 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 (G4:388). The core of morality, Kant insists, must be “pure.” But while the pure moral law would apply to other rational beings as well as to human beings, this moral law is nonetheless a central part of human beings qua rational beings, and Kant’s particular applications of the moral law increasingly emphasize the nature of human being in particular. Moreover, Kant’s dismissal of anthropology at the core of morals is really only a dismissal of *empirical* anthropology at that core. As in the case of cognition, Kant’s transcendental anthropology focuses on human actions “from-within” rather than empirically and in particular explores both the norms governing human 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 offers 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 actual, and by laying out “laws of freedom” that govern free human agents (G 4:387).

Kant’s argument for humans free agency is based on the nature of moral obligation. For Kant, the from-within standpoint of volition – where one seeks to discern what 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 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” (Allison 1990:5, 40) claims that from within deliberation, all 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 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). Kant’s way of dealing with the possibility that freedom is an illusion shifts between his *Groundwork* and his *Critique of Practical Reason*. In the former, he offers an independent argument for freedom based on humans’ cognition of ideas of reason, and thereby establishes our participation in an “intelligible world” and thus our susceptibility to moral norms. By the time of the *Critique of Practical Reason*, Kant rejects this argument in favor of a more straightforward regressive argument, assuming the legitimacy of moral norms in general and arguing *from* those norms to human freedom.⁴ Thus Kant turns from the generic perspective of deliberation to the more specific stance of one asking the question, “what *ought I to do?*”, where “ought” is specifically *moral*. In the *Critique of Practical Reason*, Kant insists that this moral “ought” is ever-present within human practical deliberation: “we become immediately conscious [of the moral law] (as soon as we draw up maxims of the will for ourselves)” (5:29). In the process of devising and considering principles to act on, we become aware of a “fact of reason” (5:29), the fact that we are bound by a moral law, which commands obedience regardless of other incentives.

From this fact of reason, Kant aims to establish that human beings are 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). In order to establish this mutual implication, Kant draws on “common rational moral cognition” to “search for and establish 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 such things. Kant puts this point in terms of a distinction between what he calls “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). Another person *who is relevantly similar* to me (able to swim, or possessed of similar talents) will have the same

⁴ For discussion of this shift, see Ameriks 2003:161-92.

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.

[S]ince 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.

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 act for the sake of ends (4:427). 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: “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, see too 6:462). 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.

Precisely what Kant means by that “human being” which is a necessary end is hotly contested, with proposals ranging from Christine Korsgaard’s suggestion that any being with a mere capacity for making choices is a human being (Korsgaard 1996:17) to Richard Dean’s recent suggestion that only a person with a wholly good will counts as a human being (Dean 2006:8). The argument for the importance of humanity in the *Groundwork* starts from the fact that “the human being necessarily represents his own existence” as an end in itself, emphasizing that this is merely “a *subjective* principle of human actions” (4:429). But, Kant argues, since all other people have the same grounds as oneself for ascribing value to themselves, one ought to regard them as ends in themselves as well. The question, then, is what any human chooser necessarily but subjectively values in making choice, and Korsgaard’s view that what is always valued in these cases is precisely the capacity for choice that is exercised in making any choice

makes the best sense of Kant's position. Choosing in accordance with the moral law requires respecting the capacity for choice of other "human beings," that is, other choosers. In this context, then, "humanity" for Kant is not limited to specifically *human* choosers, but to any rational choosers. In that sense, we must respect the "humanity" of angels, extra-terrestrials, gods, or other rational animals, if there are any such beings. But even without implying that it is *uniquely* human, Kant picks out rational agency as particularly central feature for human beings and places it at the center of his moral theory.

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: "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 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 rather than grounds morality). 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).

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 a 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 and 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)

Within deliberation, when considering whether to act on the basis of the moral law, one sees it 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

to which one can merely “give in.” 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* transcendental freedom, so a will under the moral law is transcendently free. Kant thus decries any traditional form of compatibilism as “wretched subterfuge,” mere “psychological or comparative” freedom “no ... better than the freedom of a turnspit” (5:96-7).

To his abstract argument and strong 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 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 (lust) when the fulfillment of these threatens more important sensuous desires (love of life). But the second part shows that human beings recognize in themselves an ability to overcome even 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.

Importantly, Kant is not denying that one can observe empirical causes for actions. Even consciousness of the moral law appears as an empirical cause in a chain of mental events that gives rise to a volition to act in accordance with it. But from-within volition, we become aware of a sense of responsibility the condition of possibility of which is the transcendental freedom that, properly understood, Kant’s *Critique of Pure Reason* showed to be compatible with causal-determinist explanations from-without. Through examining the moral law present within 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. Humans’ sense of moral obligation, properly understood, provides evidence of freedom and also gives rise to a specific principle of morality.

Kant’s arguments for transcendental freedom as central to human nature are hardly beyond controversy, and future chapters will take up various objections to it, but the rest of this section focuses on two problems for the Kantian account of freedom and morality offered so far.

First, 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* (see, e.g. Sidgwick 1901 and Reinhold 2006). If the moral law is the law of freedom, then whenever human beings act contrarily to the moral law, they must not really be free. But freedom is a condition of possibility of moral responsibility, so whenever human beings act wrongly, they are seemingly not morally responsible for their actions. Kant does claim that human beings can be held responsible for acting badly, but how can he do this? Second, 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 to 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 too abstract to provide real guidance for action.

Regarding the first problem, 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 *both* free beings subject to the moral law *and* subject to empirically-informed desires and inclinations. Even Kant's "pure moral philosophy" articulates what morality means for beings like us, who participate in both an intelligible world governed by laws of freedom and a sensible world governed by laws of nature. Thus while *Groundwork* begins with the "good will" in general, Kant quickly specifies the nature of this will such that it applies more particularly to wills "under certain subjective limitations and hindrances" (4:397). Because of our sensible nature, human beings have natural inclinations that can conflict with moral demands in particular circumstances. Because we have such non-moral inclinations, morality takes the form, for humans (unlike for God), of "duties" and "imperatives," commands that we *ought* to obey rather than a moral law that we simply *do* obey (4:413).

In that context, Kant distinguishes between "positive" and "negative" freedom. 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 (G 4:446-7, MM 6:213-14). Negative freedom is necessary in order to hold human beings morally *responsible*, while positive freedom constitutes the full-blown autonomy of a morally *good* agent. But there is 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 law the law of freedom, the categorical imperative. As one commentator has put it, "by making the [categorical imperative] its principle, the free will retains the position of [freedom]" (Korsgaard 1996a:166, cf. 6:227). By contrast, "the free will that puts inclination above morality sacrifices its freedom for nothing" (Korsgaard 1996a:167). 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 autonomy in the face of temptation.

The second problem, as Hegel classically put it, accuses Kant's categorical imperative of being an "empty formalism," an "abstract universality, whose determination is ... without content" (Hegel 1991:162). Precisely because this "specific principle of morality" is purely formal, it gives only the most abstract account 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. 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*” (Hegel 1991:162). Kant’s abstract moral law seems insufficient to provide moral content from within.

Kant has a two-fold response to this objection. First, even if the categorical imperative is, in itself, formal, it is still action-guiding. Hegel suggests that the categorical imperative would only prohibit stealing, 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 child’s “I will take this (other child’s) crayon to draw a picture” does not directly violate the categorical imperative because while this maxim implies crayons’ value for drawing, it does not directly commit the child to institutions of private property. 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 *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 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 free rational 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 (6:217). The result, when Kant turns to his *Metaphysics of Morals*, is a detailed account of human obligations in the face of our finite natures, both a “doctrine of right” laying out the rules governing human actions in the context of an empirical world where conflict is possible and a “doctrine of virtue” laying out particular ends human beings need to pursue given our particular predispositions, talents, and needs.

In the end, Kant’s transcendental anthropology of desire offers a detailed answer to the question “What ought I to 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 within the standpoint of choice, we can justifiably believe that humans are transcendently free things-in-themselves, even though we can never strictly “know” this. 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.

Before closing this section, it is worth attending to one further, dramatic 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 when Kant considers what the ultimate goal 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). 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 complete highest good.

Kant's arguments for God and immortality are 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 *The Metaphysical Foundations of Natural Science* makes Newtonian physics a subset of a transcendental anthropology of cognition, Kant 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 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 defended even God's existence as part of philosophical *anthropology*.

IV. What may I hope?

The Critique of Judgment as transcendental anthropology of feeling

Given the results of the previous two sections, Kant's transcendental anthropology might seem 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. 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 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 this anthropology.

First, Kant came to see feeling as capable of a priori, transcendental investigation. Kant saw the pleasures humans take in what is merely “agreeable” – food, sex, reputation – as empirically-rooted and thus incapable of a priori investigation. But as he continued to teach and study aesthetics, he came to see that judgments about beauty 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. Normativity, perhaps even of an a priori sort, is applicable to aesthetic 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” (5:196, cf. 5:176). These 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 unrelated.⁵ 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*. While the “Critique of Aesthetic Judgment” lays out an a priori principle governing feeling, there is no direct reference to feeling in the “Critique of Teleological Judgment.” And while the “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?,” Kant’s aesthetics seems irrelevant to that question. All of this can make it seem that however helpful this book might be in other respects, it cannot provide the 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. As transcendental anthropology, this a priori principle of the *human* power of judgment provides a basis for universal norms governing feeling. Purposiveness emerges as an a priori

⁵ For a defense of the claim that they are not significantly related, see MacFarland 1974. For a detailed reading of the *Critique of Judgment* as a coherent whole, see Zuckert 2007.

principle for aesthetic feeling (20:244), but once established as an a priori principle, Kant use purposiveness to supplement insufficiencies in his transcendental accounts of cognition and volition. What starts in a transcendental anthropology of (aesthetic) feeling becomes the unifying principle of Kant's transcendental anthropology as a whole and a partial basis for answering the final question of Kant's philosophy: "What may I hope?"⁶

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. 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. 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, it does not seem *a priori*.

Kant raises many of these concerns himself. He points out 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. There is also 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 volition. The necessity 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; the gratification we find in them "arouses inclination" (5:207). Useful objects are the necessary or helpful means to some given ends, so they provide satisfaction "only as a means" (5:207). And the morally good is the object of the categorical imperative; we feel satisfaction in the morally good because of its connection to volition.

But Kant claims that some pleasures are due to neither agreeableness nor goodness in their 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 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*

⁶ Ultimately, answering this question depends upon Kant's philosophy of religion and history as well (see 11:429), but the *Critique of Judgment* provides an important starting point.

(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 arousing volition). The second claim brings up the 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). Aesthetic judgments involve “a feeling of pleasure ... which ... is nevertheless ... expected of everyone” (5:191). Like other key claims in Kant's transcendental anthropology, this insistence that human beings take pleasure in the object is not an empirical-psychological claim; one does not claim that all others *will* or *do* feel this pleasure, but rather that they *should*. Aesthetic judgments present a normativity that is reducible neither to epistemic norms (since epistemic norms apply to the formation and application of concepts) nor moral-practical norms (both because aesthetics is disinterested and because practical norms require an appeal to concepts).

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” (see 5:220). In explaining what this means and how the purposiveness of beautiful objects can ground universal pleasure, Kant 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 human 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*.

The *Critique of Judgment* deals with an incompleteness in Kant's transcendental anthropology of cognition that arises because while the *Critique of Pure Reason* showed that changes in the world must happen according to causal laws, 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 (5:183, but cf. A642-68/B670-96). But human beings seek just such systematic interconnections, so 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, we 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, and the 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). 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. The *Critique of Judgment* thus fills in an important gap in the *Critique of Pure Reason*.

But Kant goes on to connect the purposiveness of nature for cognition with humans’ faculty of feeling pleasure. The connection is, at first, fairly straightforward: for human beings, “the attainment of every end is combined with the feeling of pleasure,” so if reflective judgment gives an a priori aim valid for every human being, “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 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 of which is a necessary and universal basis of pleasure for human beings. The presumption of purposiveness in nature grounds a necessary pleasure in actually discovering such purposiveness.

This pleasure is not aesthetic because it is both interested and conceptual, and in principle, objects in the world might be purposive only in that they possess a general conduciveness to be understood. But Kant uses the dilemma of aesthetic feeling to show the nature and necessity of aesthetic representations of purposiveness. Along with his contemporaries, Kant takes for granted that there *are* normative judgments of feeling (“good taste”). Kant’s transcendental anthropology looks for conditions of possibility of such judgments. They are problematic, he argues, because they must be both subjective and universal. But given 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*. In the form of the beautiful, one’s cognitive powers are “in free play,” in contrast 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. For the feeling of beauty, the relevant cognitive powers are the imagination and understanding; when these powers play freely together, one feels aesthetic pleasure.

While there is substantial disagreement amongst commentators about the nature of this free play,⁷ 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 ... 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 any determinate knowledge. Unlike 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, 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 not towards *goals* of our cognitive powers (increasing knowledge of the empirical world) but towards the *activity* of those powers. For Kant, 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, not dependent upon 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, an aesthetic judgment carries universality. (Of course, one might still get aesthetic judgments *wrong*. One’s pleasure in an object might only *seem* to be due to

⁷ Cf. Allison 2001, Ginsborg 1997, Guyer 1979, and Zuckert 2007.

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 is 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 which one judges objects as beautiful and hence worthy of pleasure. 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).

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 presents the entirety of human mental life – cognition, volition, *and feeling* – as 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 uses his account of beauty to bridge 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 regulates the investigation of nature, 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, but he added 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. Without going into the details, the account of the sublime completes Kant’s account of aesthetic pleasure and supplements his treatment of the beautiful in three important ways. First, because 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 humans’ experience of the sublime is governed by such norms, it must be discussed. Second, the sublime provides an important balance to the contribution of the beautiful to Kant’s anthropology of cognition. 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 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 fitness 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 the same sort of respect for oneself that is constitutive of moral motivation. Kant claims, “true sublimity must be sought only in the mind of the one who judges, not in the 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 caused by the superiority of one’s humanity over the sensible, natural world. While this disharmony initially provokes displeasure, the *source* of the disharmony – one’s transcendent reason – inspires an ambivalent, but nonetheless intense and pleasurable, feeling of self-esteem. Thus while 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.

The rest of the *Critique of Judgment* does not directly address humans’ faculties of feeling. But it extends Kant’s account of judgment into biology in ways that are important for understanding Kant’s account of human beings. In particular, Kant argues for what he calls an “objective purposiveness” in nature, according to which the “natural laws” under which we subsume given phenomena (organized beings) depend upon thinking of causes of those phenomena as *for the purpose* of their effect. Because one can make sense of an organized being only as “a thing ... [that] is cause and effect of itself” (5:370), such a being is a “natural end.” When one understands the motion of a 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 in terms of 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 purposively; 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 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, while “indispensably necessary,” these principles are purely “regulative,” mere heuristics “for guiding research into objects of this kind” (5:376). For Kant’s transcendental anthropology, the addition of these teleological principles has two important implications. First, it allows a limit to the causal explanation that the first *Critique* justified. While Kant insists that *in principle* everything in nature is explicable in terms of efficient causes and even that we are required to explain nature mechanically – that is, in terms of basic properties of matter – as much as possible (5:379, 429), he concedes that for humans studying the living world, such 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 living things in accordance with a principle of purposiveness gives rise to two further implications that will prove important for Kant's anthropology as a whole. The first is 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, humans require teleological explanations of some basic biological powers. 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 introduces "the idea of the whole of nature as a system in accordance with the rule of ends" (5:378-9). This yields fruit in a scientific ecology studying organisms' interdependence, and Kant insists that such study naturally leads one to think about what could be the "final end" of nature as whole.

In his transcendental anthropology of volition, Kant has already shown that humanity is a final end-in-itself. But this end-in-itself requires a transcendental freedom that cannot be an end *of nature*. Once we know that human beings as transcendently free choosers are *final* ends, however, we can look for an "*ultimate*" end, "that which nature is capable of doing in order to prepare [the human being] 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. For Kant, this *ultimate* end is humans' "aptitude[s] for setting ends at all and ... using nature as a means appropriate to the maxims of free ends in general" (5:431). 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 three. 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.

Kant's *Critique of Judgment* is a 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. 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 the necessary role of purposiveness in regulating humans' study of living things. These human principles of judgment 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 find) evidence that nature as a whole cooperates with our highest moral vocation.

Summary

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 (freely) imposes categories to cognize objects. Human action involves subordinating subjective and therefore finite maxims to an autonomous (free) 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 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 phenomena, they must be understood in terms of categories of the understanding and forms of intuition. But even as homo phenomena, humans are still distinct from merely physical nature in that we are teleologically-ordered biological organisms with particular features, many of which 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.

The transcendental anthropology in Kant's critical works not only sets up the general framework of phenomenal-noumenal humanity but also specifically addresses the *noncausal* laws that govern human beings, providing a normative account of the human being from-within: an epistemology, an ethics, and an aesthetics. However, "transcendental anthropology" is incomplete as an overall answer to the question "What is the human being?" 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 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). Finally, while the practical postulates of God and immortality and the general teleology revealed through beauty and biology give *some* basis for moral hope, "experience and history" provide further reasons that "we should not despair about our species' progress toward the better" (7:329). While developing his *transcendental* anthropology, then, Kant also pursued *empirical* studies of human beings, to further answer the question "What is the human being?" and thereby better 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.

