The Preface(s) to the *Critique of Pure Reason*

It doesn’t take long in reading the *Critique* before we are faced with interpretive challenges. Consider the very first sentence in the A edition:

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

Where to begin? We will have to say something about this thing Kant calls “human reason,” but the first thing I want to look it is this rather strange (at least for us) word he uses, “cognitions?” What exactly are *cognitions*?

The German word that this edition translates as “cognition” is “erkenntnis.” Roughly speaking, “erkenntnis” means “knowledge,” and the previous standard translation (by Norman Kemp-Smith), that’s how “erkenntnis” was translated. The problem however (or, better, the first sign that there might be a problem) is that Kant also uses the plural, “erkenntnisse.” But it seems strange to translate that as “knowledges,” as we don’t usually use this word in the plural. Kemp-Smith dealt with this by translating the singular, “erkenntnis,” as “knowledge,” but translated the plural as “modes of knowledge.” Now, this may seem like a small thing, but “knowledge” and “mode of knowledge” don’t obviously mean the same thing, and so it seems odd to translate the singular term one way and the plural term the other.

And it at least suggests there may be something important here. We think, when we read the text, that we have some relatively clear idea what Kant is talking about when he talks about “knowledge,” because we think we know something about what knowledge is. But the fact that he uses the term in the plural--something that doesn’t really make sense given what we think we know about what knowledge is--suggests that perhaps he is talking about something at least slightly different than what we understand by the word “knowledge.”

To complicate matters, Kant also uses the “kenntnis,” and the verb form, “kennen,” which are translated “knowledge” and “to know.” Are these words mere synonyms, or is there some subtle difference in how Kant uses them? So, how should these words be translated?
One of the differences with the current translations is that it translates “erkenntnis” as “cognition,” and “kenntis” as “knowledge.” I think this is helpful for at least a couple of reasons: First, since the word “cognition” sounds odd to our ears, this very oddness helps serve as a reminder that we have to be careful in thinking that we know exactly what Kant means by this term. Over 200 years have passed since Kant wrote the *Critique*, and just assuming, without real thought, that Kant meant by “erkenntnis” what we mean by “knowledge” might be asking for problems.

But this still leaves us with the question of what, as at least a working hypothesis, we are supposed to understand by this English word “cognition.” Now, my German is very weak, but I think there may be something to the difference in the two German words, “erkenntnis” and “kenntnis.” I don’t want to make too much of this, but I know that the German language often connects old words to make new ones. So, what does adding this “er” in front of “kenntnis” suggest? My sense (which really comes more from my reading of Kant than my understanding of the subtleties of the German language) is that if “kenntnis” means “knowledge,” “erkenntnis” means something like “it-knowledge.” The preceding “er” suggests that we are not talking about knowledge in general (what today we might call propositional knowledge), but knowledge of an “it,” knowledge of a thing.

Whatever the etymology of the word, I think that “cognition” should be understood as something like “knowledge of an object.” Of course, even that expression sounds odd in the plural, so perhaps a better (though wordier) phrase would be “veridical experience of an object,” i.e., an experience of an object which, because it is veridical, can be called a knowledge of an object. I think a cognition is a certain kind of mental state--those mental states which constitute “knowings” of an object.

The real point here, I think, is that “cognition” is one of those words that we need to remain somewhat open about. In the end, “cognition” is simply the totality of whatever Kant says about it. Clearly, it has something to do with knowledge. But we need to keep looking at what he actually says, rather than simply take for granted that we know exactly what *erkenntnis* is supposed to be. And so I think this odd word, “cognition,” may help us do just that.

On to “reason!”
Since this book is called a Critique of Pure Reason, some care about this word is called for as well. Kant tends to speak of their being three “faculties” of knowledge—Reason, Understanding, and Sensibility. And though this may not help much, we can probably understand a “faculty” in this context as a “source.” So, there are three “sources” of knowledge, and Reason is one of them. Even this is fraught with questions. We might say that there are three kinds of knowledge, and this is on the right track, but there are other suggestions that that these three kinds of knowledge, though distinct, are at some other level part of a “deeper unity.” (Anyone see “three persons in one God” floating around in the background?) In some sense, reason is held to be the most “fundamental” of these three faculties of knowledge.

Whatever a “faculty” of knowledge is supposed to be, Reason is the domain of general principles. The second paragraph of the Preface (immediately following the citation above) reads:

Reason falls into this perplexity through no fault of its own. It begins from principles whose use is unavoidable in the course of experience and at the same time sufficiently warranted by it. With these principles it rises (as its nature also requires) ever higher, to more remote conditions. But since it becomes aware in this way that its business must always remain incomplete because the questions never cease, reason sees itself necessitated to take refuge in principles that overstep all possible use in experience …. But it thereby falls into obscurity and contradictions …. The battlefield of these endless controversies is called metaphysics.

The first thing I want to point to here is that Reason, by “its nature” always “rises” to “higher” an “more remote conditions.” I think the point here is that reasons, by their nature, always tend towards greater and greater generality. Consider the following silly example: It’s wrong for me to kick you in the shin. Why? Because it’s wrong to injure other human beings. Why? Because it’s wrong to cause unnecessary pain. (I’m not suggest we can’t keep going with this, but I think you get the point.) In each move here, I am looking for some explanation, some “reason,” to justify the previous claim. The first answer given (“its wrong to injure human beings”) explains the first claim by appealing to a more general principle. (Kicking your shin is injuring a human being, and that’s why I shouldn’t do it.) The second answer gives an even
more general explanation of the principle offered as a justification for the first principle. (Injuring humans is wrong because it is an instance of a more general prohibition against causing unnecessary pain.) It is the very nature of reason to seek the most general explanation (or reason) possible. Could I have continued the game and asked why I shouldn’t cause unnecessary pain? Of course. And I could have questioned that response, and the next one, and so on. The “business” of reason “must always remain incomplete, because the questions never cease ....” This tells us much about the “business” or reason: it seeks generality, and can never be satisfied with anything less than complete generality. And it’s not clear that any such goal is, even in principle, ever achievable. (A quick aside: if it helps, think of the search in physics and cosmology for the “Grand Unified Theory,” a perfectly general explanation of all physical phenomena.)

The passage is also helpful in understanding, at a very general level, the problem Kant sees himself as facing: in the (natural, in fact, unavoidable) attempt to give more and more general explanations of things, human reason has found itself appealing to principles so general that they transcend the bounds of human experience, and, consequently, can never be disconfirmed by experience. This might not be a problem in itself were it not for the fact that there exists an “embarrassment of riches,” i.e., too many such explanations, not only differing from one another, but in fact mutually inconsistent (“obscurity and contradictions”). Reason seems bound, by its nature, to appeal to principles that transcend experience, and so cannot by disconfirmed by it. And the resulting in-fighting between all these competing systems has left “the queen of all sciences” in bad repute.

Kant’s prescription: a Critique of Pure Reason, i.e., an analysis of “faculty” of “Reason” itself, in search of a “perfect unity,” i.e., a single over-riding principle or function that is present in all of its various and intricate aspects. What Kant wants to know is what, if anything, can be known purely on the basis of reason, and how this might be possible (what the nature of “Reason” must be in order to make this possible). And important clue is found in the second edition of the Preface, namely that “reason has insight only into what it itself produces according to its own designs....” And part of the consequence of this discovery will be that the discipline of Metaphysics must content itself to uncovering synthetic a priori knowledge that stems from the
conditions of the possibility of experience, but give up any pretense of having insight to 
questions that transcend the “bounds of sense.” (The term “Bounds of Sense” is not from Kant, 
but the title of an influential book by Peter Strawson about the Critique.)

Much of the rest of the A Preface consists in some colorful descriptions of the various 
kinds of philosophical responses to this mess that metaphysics finds itself in. But in the B 
Preface, we get more indication of the general outlines of Kant’s solution. He begins the second 
preface by wondering how it is that a branch of knowledge can enter upon “the secure path of a 
science.” Logic, we are told, has been able to accomplish because of its “limitations.” There is 
an important hint in this: Reason (and so, Metaphysics) has a domain of its own, and within that 
domain, there is much to be accomplished. But that domain is more limited than has been 
recognized, and this is the source of its current problems.

This is only a very general sort of “hint” at his strategy, so let’s keep moving and look at 
what Kant says about mathematics. This discussion begins with a bit of historical fiction about 
how it is that mathematics found its way to this “secure path of science.” He attributes this 
“revolution” to “the happy inspiration of a single man,” but his real point is something about the 
nature of this “transformation.” Kant claims:

A new light broke upon the first person who demonstrated the isosceles triangle 
…. For he found that what he had to do was not to trace what he saw in this 
figure, or even trace its mere concept, and read off, as it were, from the properties 
of that figure; but rather he had to produce the latter from what he himself had thought into the object and presented (through construction) according to a priori concepts, and that in order to know something securely a priori he to ascribe to 
the thing nothing except what followed necessarily from what he had himself put into it in accordance with its concept. [Bxii]

This passage is still only a “hint,” so I don’t want to spend too long on it, but there are some 
interesting things here. The first point is that what we know about triangles does not arise from 
any kind of inductive inference from what we actually see in the triangle. We don’t say 
something like, “this triangle is this way, that triangle is this way, etc., so all triangles must be 
this way.” We don’t simply look to experience to tell us what the properties of a triangle are. 
But we also do more than simply analyze the definition of a triangle in order to “read off” the
properties it must have. This “revolution” comes from realizing that what we know a priori about triangles comes neither from empirical evidence nor (merely) from conceptual analysis. (How Kant thinks this knowledge is possible is hinted at here, but is not fully discussed until the Aesthetic.) The point, again, is that a priori knowledge is possible only where, and to the extent that, there are properties that we “think into the object.” Again, this is only a hint, and highly metaphorical. But a priori knowledge arises out of some kind of “contribution” from us.

The discussion of the natural sciences is equally cryptic and fanciful, and is again presented as a kind of historical description (“When Galileo fist rolled balls … a light dawned on all who study nature.”). But the point, again, is to suggest that progress in this area is the result of recognizing the role of reason. The conclusion is similar: that “reason has insight only into what it itself produces according to its own design,” i.e., into what “reason puts into nature.”

Now, what it means for reason to “put” something “into nature” is a matter that goes to the crux of what transcendental idealism really means. This is something I will come back to. But, knowing a little bit about how Kant’s position is actually developed in the course of the Critique, maybe I can say a little bit here about what he is trying to say with all this talk about “leading-strings” and “accidental observations.”

Unlike mathematics or logic (the two previous examples), the natural sciences obviously require observations based upon sense experience. But, Kant is suggesting, they require more than this. Science does not make progress simply by accumulating more and more data gathered from sense experience. When we make observations and conduct experiments, we do so to answer questions that we have already formulated. Science makes progress because we go out looking for evidence about what “laws” govern nature. But that there are laws there to be found is not something that any amount of experience could ever establish or even suggest. Without this assumption—that the natural world is bound by necessary laws—“science” could never amount to more than a “random groping” amid a patternless collection of “accidental observations.” So, what reason “puts into nature,” as Kant will explain in more detail later, is that the world that we observe (or the world as we observe it) is law-like: what reason puts into nature are laws.
But all of this explanation presupposes arguments and explanations that Kant does not offer until later in this work. At this point, all that he has really established is that the discipline of metaphysics is a mess, and that we might learn something about how to fix this mess by looking at disciplines in which progress has made. We are looking to know what, if anything can be known by “pure reason.” The key theme in all of these examples is that what we can know by reason alone are only things that are somehow “contributed” by reason. What it might mean for reason to “contribute” something, what it might mean for reason to “put” something “into nature” is the next topic for us to talk about.

Metaphysics, again, is the discipline that purports to discover “substantive” truths about reality on the basis of reason alone, without the help of (or the possibility of disconfirmation by) sense experience. This hasn’t been working so well, it seems, and Kant is suggesting that he has a way of fixing all of this, but one that involves some kind of fundamental change in the way we view the relation between knowledge and the objects known. Understanding the possibility of a priori knowledge of the world (i.e., of metaphysics) has been impossible given the assumption that “all our cognition must conform to the objects.” But, Kant thinks, we may have better luck if we assume instead that “the objects must conform to our cognition.” This is Kant’s “Copernican Revolution,” which will later (when the details are spelled out) be called “transcendental idealism.” Understanding what Kant says and evaluating why he says so, will be the focus of the rest of the course. But what is interesting here, and in the passages that follow, is text both suggesting, and then undermining, the classical interpretation of what Kant’s transcendental idealism really asserts. So what is this Copernican Revolution that is being described here?

Transcendental idealism, again, is the claim that we know only appearances and not things in themselves. (So, the “Copernican Revolution” described here is that metaphysics has floundered on the assumption that we know things in themselves, but will fare better on Kant’s claim that we know appearances instead.) On the conventional understanding of these terms, things in themselves and appearances are numerically distinct. Things in themselves are things that exist, and have whatever properties they have, independently of our experience of them. They are metaphysically real, and exist “outside” the mind. Appearances are not metaphysically
independent of us, but are in some sense “constructed” by the mind. As Kant says at one point: they are “mere representations in us.” So on this understanding of transcendental idealism, Kant is saying that the objects of experience are not metaphysically real things in themselves, but mere appearances that exist only in our experience.

The claim that we ought to reject the claim that “cognition must conform to objects” in favor of the claim that “objects must conform to knowledge” certainly seems consistent with this understanding of transcendental idealism. The idea that “cognition must conform to objects” (i.e., the view that has hindered progress in metaphysics) suggests a picture in which the objects of experience exist independently of our experience of them, and that we achieve cognition of them when our mental states conform (or correspond) to the nature of these independent objects. On this view, Humean skepticism concerning metaphysics seems justified: if these objects in no way depend upon our knowledge of them, how could we possibly know anything about them except through sense experience? Why should these objects conform to the ways our “Reason” seems to demand? Metaphysics, indeed, seems doomed.

On the other hand, if the objects of experience are mere appearances in us (object that “conform to our cognition”), the possibility of a priori cognition of them seems far less problematic. Since these appearances (in some way) depend upon our experiences of them, it is easier to understand how reason could “put” something “into them.” As mental constructs, they are constituted by our experiences of them, and so having a priori knowledge of non-trivial truths about them seems far less problematic than on the alternative view.

I have already described my philosophical reservations with this kind of reading of transcendental idealism: this account of experience seems open to attacks from both Hume and Berkeley, and so apparently justifies the complaint that Kant’s transcendental idealism is just warmed-over Berkeleian idealism inconsistently coupled with a “pre-critical” commitment to the very things in themselves that Berkeley’s view explicitly denied. If this is Kant’s view, then any lasting value that the Critique might have could only come from disentangling all this metaphysical nonsense from the remainder of the text. And, as I said, this passage (and many others throughout the text) certainly seem to support this reading.
If my only reasons for question this reading of Kant were because I didn’t like the philosophical position it attributed to him, then I suppose the proper thing to do would be to stop bothering with Kant and write my own book on the subject. I must confess that my qualms with this whole “family” of theories of mind fundamentally influenced my initial reading of Kant. But whatever my motivations for looking for them, there are real textual tensions for this understanding of transcendental idealism. One of the clearest instances of this tension occurs only a few pages later in this edition of the Preface, where Kant extols his position as having made a “distinction between things as objects of experience and the very same things as things in themselves ….” [Bxxvii] (Another such passage occurs at B69 where Kant claims that “…this object as appearance is to be distinguished from itself as object in itself.”) Such claims explicitly deny that things in themselves and appearances are numerically distinct, and so directly undermine the “two object” interpretation of transcendental idealism.

By themselves, these two passage might simply be written off as due to sloppy writing or inconsistent vestiges of a world view Kant has not yet quite given up. But the passages also suggest an alternative understanding of how to understand all this talk about things in themselves and appearances. These passages suggest that what Kant is claiming is that we do not know things as they are in themselves, but only as they appear to us. This sounds like a claim not about what things we know (mind-dependent appearances versus metaphysically real things in themselves), but instead as a claim about how we know things: because consciousness, as such, always includes an active, interpretive element, we can know and experience objects (any objects, our inner mental states just as much as objects outside us in space) only as they appear to us “through” consciousness, but never as they are independently of our consciousness of them.

I will have much more to say about this at various points in the text. For now, let me note that this notion that there is only one object, considered in two ways, seems the clear implication of the remarks about free will and the soul that follows the citation above. And if you now go back and re-read the parts of the text that introduce Kant’s Copernican Revolution--where the “two object” reading seemed so obvious, I think you will find the text far more ambiguous and interesting. I’m not going to review all of that here, but you will hear me making similar claims again and again throughout this course.