

# 11

## 'HOW IS METAPHYSICS POSSIBLE?'

### Kant's Great Question and His Great Answer

*Nicholas F. Stang*

#### 1. Introduction

There are at least two ways to be a 'great' philosopher. The first is to ask a *question*, either a question that has never been asked before, or one that has never been asked in precisely that way, or perhaps even to revive an old question that has been forgotten. The second is to give an *answer* to a question, one that is more correct, more interesting, or more plausible—or whatever other virtues that answers to philosophical questions can have—than previous answers to that question. Of course, it is possible to do one without the other. Heidegger asked a question ('What is being?') that had been asked many times before in the history of philosophy, but which had, according to him, been forgotten. One can think that, in doing so, Heidegger was doing something 'great' in philosophy even if one thinks that the various answers he gave to that question over his lifetime were incorrect. Conversely, Frege gave a new answer to a question ('What are numbers?') that is almost as old as philosophy, an answer that is, arguably, better than any answer that had been given before.

Kant has the rare distinction of both asking new questions and offering systematic, plausible, and well-motivated answers to them. Rare, but not unique: Plato and Aristotle spring immediately to mind as philosophers who did both. Kant has the further distinction (also shared by Plato and Aristotle) of having done so in virtually every one of (what we now think of as) the main areas of philosophy: logic, metaphysics, epistemology, philosophy of mind, philosophy of science, ethics, political philosophy, and aesthetics. While I think there is a way of seeing those questions, and those answers, as specifications of a single question and a single answer ('To what does reason answer?' 'Itself:'), I will focus, for the sake of space and because another chapter in this volume addresses Kant's moral philosophy, on Kant's achievement in metaphysics.

While some of his early works from the 1750s and 1760s are outstanding works of philosophy and repay close attention, his claim to greatness rests on the *Critique of Pure Reason* (*CPR*: 1st edition 1781; 2nd edition 1787), and the ‘Critical’ system of philosophy he constructed on its basis. This is a verdict shared by Kant himself, who wrote that ‘prior to the development of critical philosophy there had been no philosophy at all’ (Ak. 6:206).<sup>1</sup>

Kant asks, and answers, many questions about metaphysics in the *CPR* but they are all downstream of a single guiding question, a question he famously formulates in the Introduction to the second edition as ‘How is metaphysics possible as a science?’ (B22). The ‘possibility of metaphysics’, however, can refer to several different things and, consequently, Kant’s main question contains at least three separate questions about the possibility of metaphysics:

1. How is metaphysics possible *epistemically*? How is it possible for us to attain *knowledge* (or whatever epistemic state we are aiming for) in metaphysics?
2. How is metaphysics possible *scientifically*? How is it possible for us to attain in metaphysics not only knowledge, but the systematically organized knowledge Kant calls ‘science’ (*Wissenschaft*)?
3. How is metaphysics possible *semantically*? How is it possible for our claims to have the content, reference, and other semantic properties that they must have in metaphysics?

Each of these questions corresponds to a distinct way in which metaphysics might be defective: it might be epistemically defective because we can never acquire knowledge in metaphysics; it might be defective as a science because whatever knowledge we have of metaphysics can never rise to the level of a science; and it might be semantically defective because our claims can never have a properly metaphysical content or our concepts can never refer to properly metaphysical topics, and so forth.

Kant gives systematic answers to all three questions in the *CPR*, but in a single chapter I cannot hope to address all of them. Much contemporary Kant scholarship has focused on Kant’s answer to (1) and the related problem of synthetic *a priori* knowledge. Instead, I will focus on Kant’s answer to (3) and his account of the semantic possibility of metaphysics.

I do so for two reasons. First, I think that his semantic critique of pre-Kantian metaphysics is more fundamental to his overall argument than his epistemic critique is. Second, I think that Kant’s question about the semantic possibility of metaphysics is philosophically deeper than the other two and poses a powerful challenge even to contemporary metaphysics in the analytic tradition.

In addition to the distinction with which I began, one can also make another distinction in the ways that philosophers can be great. A philosopher can be great in virtue of making an important intervention in a specific historical context, or in virtue of some contribution (for example, a question, or an answer) whose importance transcends its specific historical context and can continue to

be significant long after that historical moment has passed. These are by no means mutually exclusive, and I will be arguing that Kant's question about the semantic possibility of metaphysics is one such instance.

In section 2 I will argue that the core of Kant's semantic critique of metaphysics is a question about the *reference* of metaphysical concepts: assuming that concepts in metaphysics refer, what explains this putative fact? In unpacking what this question means, I try to formulate it in terms more general than Kant's specific cognitive semantics, in order to facilitate the project of the rest of the chapter, in which I argue that neither pre- nor post-Kantian philosophers have adequately addressed it.<sup>2</sup> In section 3 I explain how Kant came to this question by generalizing a question that Hume raised about the concept of cause and effect; I then show why Kant's predecessors (for example, Locke, Leibniz, Wolff, and Baumgarten) lacked the resources to explain satisfactorily the reference of metaphysical concepts. In section 4 I examine contemporary metaphysics and theories of reference, and argue that it is at least questionable whether contemporary metaphysicians are better able to answer Kant's question than pre-Kantian ones are. I argue that this 'explanatory gap' is a serious problem for such metaphysicians, for it entails that they cannot explain why metaphysics is (semantically) possible. Kant's question, then, is not only of historical interest; it also poses a deep challenge to the possibility of both pre-Kantian and contemporary metaphysics, and thus constitutes a problem that any metaphysician should grapple with. In section 5 I explain Kant's own answer to this question, while omitting (for reasons of space) his argument that this is the *only possible* such answer.

The target audience for this chapter is not Kant specialists—their conviction that Kant was at least a very good philosopher can be taken for granted—but rather philosophers, especially metaphysicians, who remain unconvinced that Kant was really all that 'great' when it comes to metaphysics. Consequently, I will focus less on defending my interpretations of Kant's texts than on motivating Kant's question and his answer.

## 2. Kant's Question

In a famous letter from 1772 to his former student Marcus Herz, Kant writes (Ak. 10:129):<sup>3</sup>

I noticed that I still lacked something essential, something which, in my long metaphysical studies, I, as well as others, had failed to consider, and which in fact constitutes the key to the whole secret of metaphysics, hitherto hidden even from itself. I asked myself this question: What is the ground of the relation [*Beziehung*] of that in us which we call 'representation' to the object?

Later in this letter Kant announces the aim of systematically answering this question in a work that he anticipates will be done within three months. In fact

he did not finish it for another nine years, and it appeared under the title *Critique of Pure Reason*. Many streams in Kant's thinking fed into the writing of the *CPR* and its famous question, 'How is metaphysics possible?'; but the aspect of that question on which I want to focus—'How is metaphysics possible semantically?'—begins with this passage in the letter to Herz.

Before we can understand why Kant's question is important and innovative, we must first understand what it means. 'What is the ground of the relation (*Beziehung*) of that in us which we call "representation" to the object?' First, 'representation' (*Vorstellung*) is Kant's general term for mental states (roughly equivalent to 'idea' in Locke or 'perception' in Hume). It becomes clear in the rest of the Herz letter, and in other texts, that Kant is particularly concerned with the relation of a specific kind of representation to their objects: namely, concepts. At this point in his philosophical trajectory, he finds the 'relation' of the other main class of representations (intuitions) to their objects relatively unproblematic.<sup>4</sup> Concepts are general representations: a concept does not represent an individual object; it represents a general class of objects (Ak. 9:91). Thus, my representation of all red objects is a concept (the concept <red>) while my representation of a particular fire truck is not.<sup>5</sup> While Kant's question is originally raised at the level of representations *überhaupt*, in this chapter I will be concerned with it only as a question about specifically conceptual representation. Unpacking Kant's syntax slightly, the conceptual version of the question becomes this: what is the ground of the relation of our concepts to their objects?

However, it is harder to determine precisely what 'relation' between concepts and objects Kant has in mind in this passage. For the sake of brevity, I will simply state my own view: the relation in question is what is now called *reference*. Kant is interested in the relation between general concepts and the entities and structures in the world they are 'about': for example, the relation between the concept <substance> and substantiality.<sup>6</sup> But some of our concepts are not like <substance>; they do not refer to anything at all. Kant's examples of such 'usurped concepts' (A84/B116) are <fate> and <fortune>, but to this list we might add <witch>, <phlogiston>, and so forth. What explains the difference between the concepts that do refer and concepts that are merely 'usurped'? Why do <water> and <substance> refer but <fortune> does not? Kant's question is thus closely related to questions about linguistic reference that were discussed in mid-twentieth-century analytic philosophy (for example, 'Why does 'water' refer to water?').<sup>7</sup> Kant's question differs from those questions about reference, though, in at least one crucial respect: he raises the question at the level of thought (concepts) rather than of language (words). However, as we will see, the Kantian question about the reference of metaphysical concepts (for example, <substance>) can also be raised in a linguistic register about the reference of terms (for example, 'substance') in metaphysical theories.

In order to make Kant's question dialectically relevant to as wide a range of views in metaphysics as possible, I want to remain as neutral as possible among

different theories of what reference is, so I will opt for a very minimal characterization: reference is contribution to truth value. Our thoughts are true or false, and in some cases their truth conditions involve things in the world. (There may also be analytic truths whose truth conditions do not involve anything in the world.<sup>8</sup>) I will say that some items in the world are referred to by a thought if those items figure in the truth conditions of the thought.<sup>9</sup> For instance, the thought *water is wet* is true if and only if water is wet: since water figures in the truth conditions of *water is wet*, that thought refers to water. Intuitively, that thought refers to water because *<water>* refers to water, but this requires building more complexity into our toy semantics. For instance, if we say that in general *Fa* is true if and only if the referent of 'a' is F, then *water is wet* is true just in case whatever *<water>* refers to is wet. This only extends the notion of reference to singular terms; to extend it to terms in the predicate position, we need to add this: *Fa* is true if and only if the referent of *a* is in the extension of the referent of F. We can then ask, of some such pair of concept and referent, why that concept refers to that referent. This detour through some rather un-Kantian semantics underlines an important point: Kant's question can be raised independently of any particular theory of reference and, most importantly, independently of his own highly controversial semantic doctrines (for example, his subject-predicate theory of judgment).<sup>10</sup>

Given the centrality of the notion of reference to the rest of this chapter, it will be helpful to be as precise as possible at the outset. We must distinguish between three semantic features of a concept (or a word in our language): (1) its meaning, (2) its reference, and (3) its extension. For instance, the meaning of *<substance>* might be given by 'that of which other things are predicated and is not predicated of anything else'. But this is separate from its reference. If there are no things that are ultimate subjects of predication (it is predicates 'all the way down'), then *<substance>* fails to refer; assuming this is not the case, *<substance>* refers to *substance*, which, for ease of exposition, we can think of either as the property of being a substance or as the *kind* to which all substances as such belong (for my purposes, it will not matter greatly which we choose). The extension of *<substance>* is all individual substances. This is not to be identified with its referent. If some substances are generated or corrupted, the referent of *<substance>* does not change (it refers to substance) but its extension does. Separating reference from extension is crucial for understanding Kant's question in the Herz letter. Kant's question is not a purely metaphysical question at the level of objects (extensions of concepts), such as 'Why are there substances at all?' or 'What makes substances substances?'; it is the semantic question about why *<substance>* refers. It thus corresponds not to a purely physical question, such as 'Why is there water?' or 'What makes a sample of water water?', but instead to a semantic question about why *<water>* refers.

The Kantian question, 'Why do our concepts refer to the objects they do?', has its contemporary counterpart in the question, 'Why do the words of our

language refer to the objects they do?’ Answering that question in general (*überhaupt*) is no easier, but also no more difficult, than giving a general theory of linguistic reference. To the contemporary philosopher, as for the eighteenth-century philosopher, there are many candidates for such a theory (descriptive theories, causal theories, direct reference theories, etc.) and the challenge is to find a model that extends to all of the cases for which one wants to account.<sup>11</sup> The challenge, of course, is that some theories (for example, causal theories) are better suited to explaining some kinds of cases (for example, names for ordinary objects) than others (for example, names of numbers). Kant’s point in the letter to Herz is that none of the theories of conceptual reference given by his predecessors (theories of how our concepts are ‘related’ to objects) is much help at all in explaining why metaphysical concepts refer. And Kant’s point holds even today, with our much more sophisticated theories of reference, as I will argue in section 4.

Recall from above the two ways in which the possibility of metaphysics might be questionable: (i) we might question the epistemic possibility of achieving knowledge of our metaphysical theories, and (ii) we might question the semantic possibility of referring to anything using our metaphysical concepts. Kant does have an argument that calls into question the *epistemic* possibility of acquiring knowledge in metaphysics: knowledge in metaphysics (he argues) must be both synthetic and *a priori*, and so we require an explanation of how we could know something *a priori* if it is not an analytic truth. However, I do not focus here on the problem of synthetic *a priori* knowledge, for three reasons. First, I am making the case for the continuing relevance of Kant in metaphysics, and it is harder to show that the problem of synthetic *a priori* knowledge remains one with which contemporary metaphysicians should be concerned, since many of them have jettisoned the analytic-synthetic distinction and few are committed to the strictly *a priori* status of metaphysics. Second, I want to make a case for the (relative) *originality* of Kant’s question, and there are good reasons to think that the problem of synthetic *a priori* knowledge was known to Kant’s rationalist predecessors (though not under that name).<sup>12</sup>

Third, and most important, the reference question is more fundamental than the problem of synthetic *a priori* knowledge. Even if this problem were solved, or dissolved, the reference problem would remain. For instance, even if metaphysics were thoroughly analytic (as many of Kant’s contemporaries held), we would still ask why these concepts that we are analyzing are about anything, rather than being empty concepts like <witch> or <phlogiston> or <fate>. Even if we abandoned the assumption that metaphysics must be strictly *a priori*, and held that metaphysics is continuous with the natural sciences, and that its epistemology is broadly abductive (resting on inference to the best explanation), we could still ask why the concepts that figure in these theories refer to anything. And even if we took a Kantian direction and held that metaphysical knowledge must be synthetic and *a priori* and had an explanation of how such knowledge is possible, we would still need a *separate* explanation of why the concepts that

figure in such synthetic *a priori* knowledge refer. The reference question is thus separate from, and more fundamental than, the knowledge question.

This raises two questions. What is so hard about explaining the reference of metaphysical concepts? And (why) must the metaphysician offer such an explanation? I address these questions in the next section.

### 3. Kant's Question Before Kant

While none of the pre-Kantian philosophers with whose writings I am familiar devotes the focused attention to the reference of metaphysical concepts that Kant does, it is not hard to find in their writings implicit or explicit answers to Kant's question. In this section I discuss several of the more prominent such answers and argue that none of them succeeds. But first I want to discuss the relation of Kant's question about metaphysical concepts to a similar question raised by Hume.

#### (i) Hume

In the Preface to the *Prolegomena* Kant writes, 'I freely admit that it was the remembrance of David Hume that awoke me from my dogmatic slumber' (Ak. 4:260). It is likely that Kant's 'remembrance' was of Hume's discussion of causation in the first *Enquiry*, specifically §VII ('Of the Idea of Necessary Connection').<sup>13</sup> In that section Hume begins by formulating a more precise version of his famous 'copy' principle: every simple idea (an idea not composed out of further ideas) is a copy of some impression, its 'original'. Hume then inquires into what the 'original' of our idea of necessary connection could be and rejects various candidates. In Part II of §VII he argues that the only suitable impression is 'a customary transition of the imagination from one object to its usual attendant'. Having observed one kind of object (for example, fire) constantly followed by another kind of object (for example, smoke), when we encounter the first object we immediately imagine the second object, and that feeling of imaginative transition is the 'original' of our idea of a necessary connection between the objects.

In the *Prolegomena* Kant writes (Ak. 4:257):

Hume started mainly from a single but important concept in metaphysics, namely, that of the *connection of cause and effect* (and of course, also its derivative concepts, of force and action, etc.), and called upon reason, which pretends to have generated this concept in her womb, to give him an account of by what right she thinks that something could be so constituted that, if it is posited, something else necessarily must thereby be posited as well; for that is what the concept of cause says.

Kant is here translating Hume's question, formulated in terms of impressions, ideas, and copies, into the terms of his own cognitive semantics. An idea, for

Hume, represents (is an idea ‘of’) more than just the individual impression from which it was copied (or the set of simple impressions from which it is copied, if the idea is complex) when it is imaginatively associated with other individual ideas that resemble the original impression in some determinate respect.<sup>14</sup> Ideas ‘used abstractly’ in this way constitute Hume’s explanation of how we represent generality. They thus correspond to Kantian concepts. Hume’s copy principle, in Kantian terms, is his theory of how concepts (ideas) ‘relate to’ (refer to) their objects (impressions). In the *Prolegomena*, and in the 1772 letter to Herz, Kant presents himself as raising Hume’s question about the idea of necessary connection, but detaches it from the copy principle and the rest of Hume’s specific cognitive semantic theory: why do concepts (general representations) refer to their objects? But Kant also generalizes Hume’s question to all of the *a priori* concepts of metaphysics (Ak. 4:260):

So I tried first whether Hume’s objection might not be presented in a general manner, and I soon found that the concept of the connection of cause and effect is far from being the only concept through which the understanding thinks connections of things *a priori*; rather, metaphysics consists wholly of such concepts.

But if Kant had read the *Treatise*,<sup>15</sup> he would have known that Hume raises this problem not only about <cause-effect> but also about many of (what Kant would consider) the other *a priori* concepts of metaphysics: possibility, existence, substance, personal identity, and so forth. Does this compromise the originality of Kant’s question? Yes, somewhat, but it does not eliminate it. For what Kant saw was that a version of ‘Hume’s problem’ arises even if one is not as radical an empiricist as Hume, and that contemporary non-Humean metaphysicians lacked an answer to this problem (see below). What is more, he not only saw that Hume’s problem generalizes (even on non-Humean premises); he saw ‘how’ to generalize it: it applies to all the *a priori* concepts of metaphysics. Rather than treating this as a piece of Kantian technical terminology to be explicated in terms of other technical terminology, I will instead consider the role that the notion of ‘*a priori* concept’ plays in the problem for metaphysics that Kant is trying to motivate. This will be easiest to see by looking at his reasons for thinking that another of his predecessors, Locke, failed to explain why metaphysical concepts refer.

## (ii) Locke

Kant writes in the *CPR* (A85/B117):

I therefore call the explanation of the way in which concepts can relate to objects *a priori* their transcendental deduction, and distinguish this from the empirical deduction, which shows how a concept is acquired through experience and reflection on it.

If the 'relation' of a concept to its object is the reference of that concept to that object (*those* objects) then a 'deduction' of a concept is an explanation of why it refers to objects.<sup>16</sup> Kant here distinguishes a 'transcendental deduction', which is said to be the only possible explanation of why *a priori* concepts (like those in metaphysics) refer, from an 'empirical deduction', which explains why a concept refers by tracing it back to the experience from which it was abstracted. Kant explains a few pages later why such an 'empirical deduction' of metaphysical concepts like *<cause-effect>* is impossible (A91/B123):

If one were to think of escaping from the toils of these investigations by saying that experience constantly offers examples of a regularity of appearances that give sufficient occasion for abstracting the concept of cause from them, and thereby at the same time thought to confirm the objective validity of such a concept, then one has not noticed that the concept of cause cannot arise in this way at all. [. . .] For this concept always requires that something A be of such a kind that something else B follows from it necessarily and in accordance with an absolutely universal rule.

An empirical deduction of a concept explains why it refers, by tracing the concept back to an experience from which the concept was originally abstracted. We can think of this explanation as having roughly the following form: one experiences an object (for example, a cause-effect pair), one forms the concept of all other objects that resemble this object in relevant respects, and this concept refers to all other such objects that resemble the object in the relevant respect. One page earlier, Kant names Locke as a thinker who attempted to give such an 'empirical deduction' of the concept of causation.<sup>17</sup>

It is worth taking the effort to understand Kant's rejection of empirical deductions of metaphysical concepts because this remains a potentially attractive strategy for contemporary metaphysicians, many of whom are not wedded to the *a priori* status of metaphysics.<sup>18</sup> Metaphysical concepts, according to this line of thought, are broadly 'empirical' concepts and their reference is explained in a fashion similar to that of other empirical concepts: one experiences an object with a certain property, one abstracts a concept of that property, and that concept then refers to any object with the relevant property. Kant rejects such an empirical deduction of *<cause-effect>* because, he claims, this is a concept of a necessary connection (the cause necessitates the effect) and experience never presents us with necessary connections. If correct, this undermines the possibility of an empirical deduction because, although we might experience events that stand in necessary connections, we do not experience them *as* necessarily connected, and so nothing about our experience explains why our concept would have this modal content. This objection brings out a key feature of a Lockean empirical deduction: my experience of an object that is F only explains why my concept refers to all objects that are F if the property F can figure in the content of my experience. On the Lockean picture, concepts get their content from the content

of experience. In Locke's own terms, the ideas that I abstract can only represent qualities that resemble sensory ideas that I perceive.

But the Lockean has at least two responses to this Kantian argument. First, one might simply *deny* that the content of experience is as modally impoverished as Kant here claims it is,<sup>19</sup> and insist (though Locke himself would not) that we simply *do* experience the necessity of the relation of cause and effect. Second, one might point out that not all concepts in metaphysics are concepts of necessary connections. Indeed, even in Kant's own theory, it is far from clear that the categories of quantity and quality (for example, <unity>, <totality>, <reality>, <limitation>) have a modal content that cannot be present in experience. So the Lockean might hold that, regardless of whether or not <cause-effect> admits of an empirical deduction, many (most) other metaphysical concepts do.

However, Kant has the resources to respond to this defense of empirical deductions, for this modified Lockean story is ultimately no explanation at all. Translating Locke's theory into Kant's terminology, the Lockean explains the content of concepts (what properties they refer to) by means of the content of experience. But if our concepts are going to refer to a certain class of properties, we need an explanation of why properties of that class are represented in experience.

Simplifying slightly Locke's complex distinction between primary and secondary qualities and focusing on the primary qualities of objects (for example, shape, size, volume), what explains the match between the properties (qualities) of the object and the content of the experience is a causal connection: some primary qualities in the object cause me to experience it as having those qualities (to perceive ideas that resemble those primary qualities).<sup>20</sup> This causal constraint allows the Lockean to explain why we are not massively mistaken in our formation of empirical concepts (Locke's abstract ideas): we do not systematically form concepts of properties that objects lack, because the presence of those properties is among the causes of the experience whose content grounds the content of the concepts. What plays the corresponding role, in the Lockean picture, in the case of metaphysical concepts? It is not enough for the Lockean to say that various metaphysical properties (causation, substantiality, unity, etc.) are represented in experience; in order to explain why concepts abstracted from such experiences refer to the properties (qualities) in objects, the Lockean must explain why those experiences, in representing such properties, do not suffer massive reference failure.

Ironically, the Lockean may be in the best position to explain this in the case on which Kant focuses: cause and effect. Arguably, the very causal powers of objects cause us to experience them as causally related. But this is not a promising strategy to pursue with other metaphysical concepts, which seem to refer to properties that are causally inert. Without substantial additional argumentation, the Lockean has given no explanation of why, for instance, the relation of inherence by which an accident is *in* a substance (Kant's first category of relation) or the totality of a set of parts (Kant's third category of quantity) plays a corresponding causal role. There are thus good reasons, both in the eighteenth and

in the twenty-first century, not to look to a Lockean 'empirical deduction' for the explanation of why metaphysical concepts refer.

Finally, we can give a provisional account of what '*a priori*' concepts are: concepts whose putative reference cannot be explained by an empirical deduction. It is because the concepts of metaphysics cannot be 'deduced' empirically that Kant seeks a different origin for them than experience, though his theory of the origin of *a priori* concepts will not be our topic here. This also shows that there are *a priori* concepts that are not specifically metaphysical concepts (e.g., mathematical concepts).

### (iii) *Logicism*

As I argue extensively in Stang (2016), Kant's predecessors in the German rationalist tradition—Leibniz, Wolff, and Baumgarten—hold what I call a 'logicist' account of possibility: it is possible that  $p$  if and only if  $p$  does not entail a contradiction. In particular, this entails that a concept is possibly instantiated if and only if the proposition that it is instantiated does not entail a contradiction. Assuming that the *a priori* concepts of metaphysics are logically consistent (an assumption Kant grants) it follows that they are possibly instantiated, and hence, in this sense, refer at least to possible objects: they refer to the possible objects that instantiate them.<sup>21</sup> For convenience, I will refer to Leibniz, Wolff, and Baumgarten as *logicists*.

Note, however, that from the fact that it is possible that there is an  $F$  it does not follow that  $F$  refers; it follows merely that *possibly*  $F$  refers. This point—that the possibility that  $F$  refers does not entail that *there is* an object (even a logically possible one) to which  $F$  refers—is closely related to Kant's famous claim that existence is not a 'real predicate', a doctrine from which (on my reading) the logicists dissented. But even if we grant them the assumption that for every logically consistent concept there are logically possible objects that fall under it (and thus that *there are* objects that may or may not exist), this shows merely that the logical possibility (consistency) of a concept establishes that it refers to logically possible objects. This answers Kant's question about the reference of metaphysical concepts only at the cost of 'lowering the bar' for what counts as reference. On this view, a 'usurped concept' such as <fortune> (or <witch>) refers just as much as <substance>: since both are logically consistent, they refer to logically possible objects. But even the logicists would want to distinguish between usurped concepts and genuinely referring concepts, for without such a distinction, metaphysics is on a par with astrology and witchcraft.

To return to a point from the previous section, this is not a request for an explanation of why there are substances; that would be an explanation of why <substance> has a non-empty extension, not why <substance> refers. In fact, by parity of reasoning the logicist explanation in the previous paragraph of why, for example, <substance> is a concept of logically possible objects is merely an

explanation of why it has a non-empty extension: there are logically possible substances. It does not explain why *<substance>* refers to substance, which for our purposes we can think of as the property possessed by all substances.<sup>22</sup> One way of appreciating this point is that logically atomic concepts are, trivially, logically consistent, but this by itself does not explain why they refer, for it does not distinguish between referential logically atomic concepts (which compose all other meaningful concepts) and logically atomic concepts that do not refer.

In his 1763 work *The Only Possible Ground for a Demonstration of the Existence of God*, Kant puts the point this way (Ak. 2:80–1; Kant 1992: 125):

Suppose that you can now no longer break up the concept of extension into simpler data in order to show that there is nothing self-contradictory in it—and you must eventually arrive at something whose possibility cannot be analyzed—then the question will be whether space and extension are empty words, or whether they signify [*bezeichnen*] something. The lack of contradiction does not decide the present issue; an empty word never signifies anything self-contradictory. If space did not exist, or if space was not at least given as a consequence through something existent, the word ‘space’ would signify [*bedeutet*] nothing at all.

Kant’s point is that the mere logical consistency of a concept does not explain why (in my terminology) it ‘refers to’ (*bezeichnet, bedeutet*) anything, because there can be logically consistent concepts that refer to nothing, as *<space>* would if space were impossible. Kant’s way of making this point is to distinguish between logical possibility and ‘real’ possibility (which we would now call ‘metaphysical’ possibility), which Kant repeatedly equates with the ‘thinkable’. In his pre-Critical metaphysical works of the 1760s, Kant argues that not all logical possibilities are really (metaphysically) possible: not all logically consistent concepts refer to really possible properties. The question then becomes that of why metaphysical concepts refer to really possible properties, and not merely to logically possible properties.

But this might seem to beg the question against the logicists, who think that logical possibility entails real (metaphysical) possibility. However, while the logicists maintain, *contra* Kant, that logical and ‘real’ (metaphysical) possibility are coextensive, they acknowledge that these are separate notions, for these have distinct definitions. A logically possible concept is one that contains no mutually contradictory marks (sub-concepts). What Kant calls ‘real’ possibility, however, they ground in the unlimited perfections of God. A logically atomic concept refers to some infinite perfection possessed by God (for example, understanding, power, will) and a logically consistent complex concept refers to some really (metaphysically) possible property composed of unlimited divine perfections. It follows that C is logically consistent if and only if it refers to some really possible property. The logicists thus seem to have a well-thought-out answer to Kant’s question about concept reference: a concept refers to really possible (not merely logically

possible) property just in case it is a logically consistent combination of atomic concepts of unlimited divine perfections.

Kant's pre-Critical works contain several complex and controversial arguments that real and logical possibility are not coextensive,<sup>23</sup> which have been extensively critically discussed by scholars.<sup>24</sup> However, I think that, in addition to the arguments that Kant gives explicitly, it is relatively clear, from a Kantian point of view, that the logicians have simply assumed precisely what stands in need of explanation: why do my concepts refer to the really possible properties to which they refer? Since the reference of complex concepts is explained by the reference of their atomic constituents, we can formulate this at the level of atomic concepts: why do my atomic concepts refer to unlimited divine perfections? The principal logicist answer to this question, shared by Leibniz and Baumgarten, is that God created me with a concept (or at least the capacity to come to reflective awareness of this concept) referring to that property. But then the Kantian question becomes this: In virtue of what did God do this? It is not enough, when pressed for an explanation, to simply say 'God did it'; one must say what God did, and why, in virtue of doing that, he made it the case that the semantic relation of reference obtains. If we assume that semantic facts about what concepts refer to are not brute facts (a premise accepted by all of the logicians), then there must be some fact *p* in virtue of which my concept <substance> refers to substantiality, and God created me with a concept that refers to substantiality in virtue of making it the case that *p*; what, then, is *p*? Nor will it help for the logicist to claim that concepts are individuated by their referents: if it is essential to my concept <substance> to refer to really possible substances, in virtue of what did God create me with this very concept with this content, <substance>, rather than another one (e.g., a usurped one, like <fortune>)?<sup>25</sup> It is not clear that logicians can answer this question.

Further pursuing this question within rationalist metaphysics would take us too far afield. In this section I hope to have established that Kant's question about the explanation of metaphysical concept reference poses a deep challenge to the metaphysics of his forebears. In the next section I will argue that it is by no means clear that contemporary metaphysicians are in a better position to answer it.

#### 4. Kant's Question Today

Since the question of metaphysical concept reference is, to my knowledge, even less discussed in contemporary metaphysics than it was in early modern philosophy before Kant, most of the 'answers' to Kant's question that I will consider are my own reconstructions from the contemporary literature, as are the 'Kantian' responses that I will discuss. Most contemporary theories of reference are theories about linguistic rather than conceptual reference, so in this section I will formulate Kant's question, and potential responses, in terms of the reference of linguistic terms in metaphysical theories.

### **Description**

One might think that the reference of a term like ‘substance’ can easily be explained by thinking of it as a disguised descriptive term: ‘substance’ just means ‘individual with properties, which is not itself a property of some further individual’. ‘Substance’ refers just in case the description refers. But this merely pushes the explanatory burden back one step; we need an explanation of why ‘individual’ and ‘property’ refer. This is why Kant raises his question about the categories, the basic metaphysical concepts in terms of which other metaphysical concepts are defined.<sup>26</sup>

### **Elimination and Deflation**

Some contemporary metaphysicians deny, in Kant’s terms, that metaphysical terms refer. Hofweber (2009), for instance, argues that syntactically singular terms for numbers, properties, and propositions do not have the semantic function of referring. Philosophers like Hofweber might thus be said to *eliminate* reference from metaphysics.<sup>27</sup> Other philosophers might be said to *deflate* the reference of metaphysical terms. According to the ‘easy’ approach to ontology defended by Thomasson (2015), sortal terms like ‘number’ or ‘property’ or ‘proposition’ are associated with application conditions, where it is analytic that, if the application conditions are satisfied, an object of the relevant sort exists and thus that the term refers. For instance, it is analytic (on Thomasson’s view) that if Beyoncé’s dress is red, then her dress has the property of being red and thus that ‘property’ refers. On Thomasson’s view, there is nothing more to the reference of these sortal terms than the satisfaction of their application conditions. Neither the eliminativist nor the deflationist need be bothered by Kant’s question about metaphysical reference. The eliminativist denies the explanandum while the deflationist can explain it (or explain it away) all too easily: she can simply point to the fact that the application conditions are fulfilled. Consequently, Kant’s question poses no problem for the contemporary metaphysical eliminativist or deflationist. Kant’s question poses a problem for the theorist who holds that reference is a substantive relation that partly explains why our metaphysical theories are true. As we will see below, there are several such theorists on the contemporary scene. I will refer to them as *non-deflationary realists*.<sup>28</sup>

### **Explanation by Division**

Some readers versed in contemporary philosophy may be unimpressed with Kant’s question in the first place, for they will think that it is all too easy to answer. Its appearance of difficulty, they will say, depends on a conflation of a psychological with a ‘Platonic’ (or Fregean) notion of a concept. If we distinguish between concepts as psychological entities or capacities and the meanings of

those concepts—'Concepts' in the Fregean sense—then we can see that the reference of a given term divides without remainder into three factors: (i) the fact that our term expresses a Concept, (ii) the fact that that Concept has the meaning it does, and (iii) the fact that that Concept refers to what it does.<sup>29</sup> The first fact is a semantic one and admits of a semantic explanation, if any. The second fact is partly constitutive of what it is to be that Concept: Concepts are individuated by meanings, so it is part of what it is to be a given Concept to have the meaning that it does—for example, part of what it is to be the Concept <substance> is to be a Concept of individuals with properties. Finally, there is some fact about what the Concept refers to. In the case of ordinary empirical terms, whether the Concept refers and what it refers to will depend upon contingent facts about the world. For instance, the Concept expressed by the definite description 'the sixth Chairman of the State Council of the German Democratic Republic' does not refer, and the fact that the Concept expressed by 'water' in the mouths of English-speakers on Earth refers to water depends on the contingent fact that water, rather than XYZ, is in our environment.<sup>30</sup>

But the terms of metaphysical theories are presumably not like this; they refer necessarily, so there is a purely metaphysical explanation, if there is any explanation at all, of why a Concept with a given meaning refers.<sup>31</sup> Even if we reject the assumption that metaphysical concepts refer necessarily if they refer at all, and allow that it might be contingent that, for instance, <substance> refers, the important point still stands: the explanation of (iii) will be at the level of pure metaphysics, for it will be an explanation of why a Concept with some given meaning refers. For instance, it might be that the meaning of our term 'substance' is the Concept <substance>, whose meaning is given by: *x* is a substance if and only if *x* is an individual with properties. This Concept with this meaning refers because, for purely metaphysical reasons, there must be individuals with properties.

The problem with this explanation by division is that, by giving distinct explanations of these three factors, it leaves unexplained why the very same Concepts expressed by terms in our language refer. It remains, on this 'multifactor' view, a mystery why, in forming metaphysical theories, we do not massively misfire and use terms that express Concepts that fail to refer.

Compare the three-factor explanation of metaphysical concept reference to a 'three-factor' explanation of why we have true beliefs: (i) we form beliefs (understood as psychological states) that (ii) express certain propositions that essentially have certain truth conditions, and (iii) those propositions have their truth conditions satisfied by the concrete realm. The first fact admits of a purely psychological explanation, the second fact is explained by the essence of these propositions (they are individuated partly by their truth conditions, let us suppose), and the third fact has a purely physical explanation (why the concrete realm satisfies the truth conditions constitutive of the relevant propositions). But this story by itself leaves unexplained why we have *true* beliefs: the facts that

make the propositions true play no role, within this story, in explaining why we form beliefs that express those propositions. Likewise, in the three-factor explanation of metaphysical concept reference, the fact that these metaphysical Concepts refer plays no role in explaining why there are terms in our theories that express them. Notice that this is not the case with concepts and beliefs about objects in our sensible environment: a given sensible fact (for example, the fact that there is coffee in my cup) or the fact that a given Concept refers (for example, the fact that *<coffee>* refers to that coffee) partly explains why we form a psychological belief that expresses a true proposition (the belief that there is coffee in the cup) or a term ('coffee') that expresses a referring Concept. But this explanation depends upon our sensory awareness of our environment, and in section 3 we saw some anti-Lockean reasons to be skeptical that perceptual experience can explain the reference of metaphysical concepts or the terms in metaphysical theories.

### ***No Special Problem***

Another dismissive response that a contemporary metaphysician might give to Kant's answer is that there is no special problem about metaphysics. We can raise explanatory questions about the reference of terms in any domain: natural scientific, mathematical, and so forth. If metaphysics is, in this respect, on a par with other sciences, then it would seem that Kant has failed to show that metaphysics is especially problematic. It would be nice to have an explanation of reference in metaphysics, but it would also be nice to have such an explanation in physics, mathematics, and so forth. There is no more reason, in the absence of such an explanation, to be skeptical of metaphysics than there is to be skeptical of physics or mathematics.

Recall that Kant's question is not the skeptical epistemic question: how do we know that metaphysical terms refer? It is an explanatory question: what explains the fact (assuming it is a fact) that metaphysical terms refer ('relate to their object')?<sup>32</sup> Once we recognize that Kant's question is fundamentally a demand for explanation, however, the 'no special problem for metaphysics' response is, to a great extent, dissolved. For there are explanations available of why the terms in physical theories refer. The most plausible of them rest, I take it, on notions like 'natural kind' (or on the more general notion of 'joint carving'—see below) and have something like the following form: the terms in our best scientific theories refer to the natural kinds that make them (approximately) true, if they are approximately true.<sup>33</sup> Kant's question is about the very terms used to explain why terms in natural science refer: why, for instance, does 'natural kind' refer to natural kinds? If we lack an answer to Kant's question, this does not undermine our claim to knowledge in physics. It shows, at most, that physics rests on an assumption—that its terms refer—an assumption that is explained only in the distinct science of metaphysics. So, in a certain sense, physics depends

on metaphysics. But the nature of the dependence is explanatory, not justificatory. The lack of an explanation of metaphysical reference does not ramify into skepticism about our knowledge of physical theories, nor does it undermine the 'metaphysical' explanation of why physical terms refer; it shows merely that this explanation relies on an assumption (that 'natural kind' refers) that is not itself explained, a feature shared by (presumably) all explanations. It shows, at most, that metaphysics cannot 'look after itself'—it cannot explain one of its basic assumptions (that its terms refer) and it cannot appeal to some more fundamental science, because it is, allegedly, the most (explanatorily) fundamental science.<sup>34</sup>

### **Reference Magnetism**

Finally, one might look to the idea of 'reference magnetism' for an explanation. The core idea of reference magnetism views is that certain entities (properties, concept extensions, etc.) are intrinsically more natural or 'suitable to be referred to' than others, and this is what breaks the tie among candidate referents that equally satisfy the descriptive contents of our theories. One might exploit a similar idea in explaining the reference of metaphysical terms: the world has a privileged metaphysical structure, and *ceteris paribus* our terms 'carve' the world at the joints of that structure.<sup>35</sup> However, for reasons of length, and because I criticize reference magnetism at length elsewhere, I will forgo any further discussion of this kind of explanation.

## **5. Kant's Answer**

Having explained the meaning of Kant's question, and why various pre-Kantian and contemporary answers to it are less than fully successful, I will now explore Kant's own answer, his own explanation of why metaphysical concepts refer. As in the rest of this chapter, I will not be following Kant's text especially closely, both for reasons of brevity (fully grounding my interpretation in the texts would require much more space than I have here) and because my primary aim remains making the salience of Kant's question and his own answer to it clear to contemporary audiences.

The key idea in Kant's own explanation of why metaphysical concepts refer is a deceptively simple one: something may have a different structure in itself than it has in relation to something else. For instance, a cow in itself has a certain anatomical structure: these and those organs related in such and such a way. But in relation to our practice of butchery, a cow has a quite different structure. A butcher (quite literally) carves a cow at different joints than an anatomist does: in relation to butchery, a cow divides into the edible and inedible (bone, skin, certain organs) parts, and edible parts divide into various 'cuts' (top round, sirloin, etc.) that are in no sense the natural 'joints' of the cow.<sup>36</sup> This is an especially

literal (and perhaps gruesome) example, but other examples abound. The physical elements have a certain structure, which is given by the periodic table of elements. The ‘natural joints’ among the physical elements are, for instance, the inert gases, the metals, the halogens, and so forth. But in relation to human biology the physical elements have a different structure with different joints. In relation to human biology the physical elements divide into (perhaps overlapping) classes such as those that are poisonous to humans, those that are essential constituents of our bodies, and so forth. The intuition behind the Lewis-Sider idea of naturalness and joint carving is that some ways of dividing things up are objectively more correct than others. But even when we consider some entity (cow) or system of entities (the elements) in relation to something else (butchery, human biology), we can still distinguish ways of dividing them up that are more ‘correct’ than others. For instance, in relation to butchery, dividing a cow into the front and back half is objectively incorrect; you would be doing butchery wrong if you divided the cow that way. Likewise, in relation to human biology, dividing the elements into those with odd and those with even atomic numbers is objectively incorrect; that is not a natural joint in the physical elements, even in relation to human biology.

Let us concentrate on one particular example of the difference between the structure of something in itself and its structure in relation to something else: the structure of the world in its own right (the topic of metaphysics, according to Wolff, Baumgarten, and Sider) and the structure that the world has in cognitive relation to us. Now, this distinction can be many different distinctions, depending on what cognitive relation is specified. For instance, the world might have a different structure ‘in itself’ than in relation to our beliefs about that structure, if our beliefs are mistaken about that structure. Likewise, the world might have a different structure ‘in itself’ than it has insofar as it is perceived by us. The world as perceived by us has a certain structure (spatiotemporal objects in motion with various perceptible properties) but, if current physics is accurate, the structure that it has in itself is quite different.

Kant holds that the subject matter of metaphysics should be understood, not as the nature of the world in itself and its structure, but as the structure of the world in cognitive relation to us. The relevant cognitive relation is not *belief* or *perception*; it is what Kant calls *Erkenntnis*. This, however, is a technical term in his philosophy, so it provides little independent purchase on what relation he has in mind. This is reflected in the now standard translation of *Erkenntnis* as ‘cognition’, which is itself a technical philosophical term without a very precise meaning independent of the uses to which particular philosophers put it.

Because this is not primarily an exegetical chapter, I will simply cut to the chase and state what I take to be the closest analogue in contemporary philosophy to Kantian *Erkenntnis*, namely, understanding.<sup>37</sup> Applying this back to the idea with which we started, we can distinguish between the structure that the world has in itself and its structure as *understood* by us. It is important to understand

that this is not the distinction between the structure that the world has in itself and the structure that we merely *believe* it to have. The difference arises from the difference between the state of understanding and the state of belief. These cognitive states are individuated by an internal standard. The individuating standard of belief is truth: one's belief that  $p$  is a successful belief (it satisfies the standard that makes belief what it is) when it is the case that  $p$ . What is the internal standard of successful understanding, the standard that (partly) individuates this achievement as the achievement that it is? One answer would be this: one understands  $X$  successfully when one carves  $X$  at its joints, when one understands the structure of  $X$  in itself. But this is not Kant's view about understanding. His view is that understanding has its own internal standard and structure that may be distinct from the structure of the object of understanding. Our understanding has, so to speak, its own terms, which may not be the 'terms' (joints) in the world itself. We successfully understand something when we understand it on our terms. The 'terms' of our understanding are the structure of anything understandable by us, and this is dictated, according to Kant, by the nature of our minds, *not* by the structure of the thing itself.

An example might be helpful. Consider teaching the definition of a derivative to a student. You first try to get the student to understand the notion of a converging sequence of real numbers, then you define the first derivative of a real-valued function at a point as the limit of a converging sequence of slopes. But your student cannot understand it in these terms; he can only understand it in visual-geometric terms: the derivative is a line orthogonal to a curve at a given point, and so forth. Your student has understood something, and that understanding has its own internal standard by which the student can succeed or fail in a given instance. For instance, if the student, given the curve for the function  $f(x) = x^2$ , thinks that the tangent of the curve at the origin points in the direction of the positive  $x$ -axis, he has failed to understand even according to the internal standard of his own visual-geometric understanding of what a tangent is. Likewise, the student has understood by his own standard if he realizes that the tangent is null at that point. The student has understood something about what you have taught him (the notion of a tangent) but has not understood it in its own terms: he has not understood the rigorous definition of a derivative as the limit of a converging sequence of slopes. Kant's point is that we are always in the position of this student: our minds are equipped with a set of concepts in terms of which we understand anything at all. Anything we can understand, we understand in these terms. These concepts describe the structure of any understanding we can possess, and thus they describe the structure of anything *as understood by us*.

It is not hard to see how to turn this into an explanation of why our concepts in metaphysics refer. First, we must distinguish two kinds of metaphysics—what Kant would call, in turn, 'immanent' metaphysics and 'transcendent' metaphysics (see A296/B352). The object of immanent metaphysics is the world *as*

*understandable by us*. Its object is not the world *as actually understood by us*, for that is far more determinate and partly depends upon empirical discoveries in the natural sciences. One of the things we actually understand about the world is that the earth is billions of years old, but this does not mean that a world in which the earth is much younger is not understandable by us. The subject of immanent metaphysics is, so to speak, the complete space of how the world can be understood (not what actually happens to be understood). It is not about that space ‘piecemeal’; it is about its structure, its joints: the topic of immanent metaphysics is the structure of the world as understandable by us. Kant’s view, as we have seen, is that the structure of the world as understandable by us is determined by the structure of our own intellects. So there is a direct translation between the structure of our understanding of the world and the structure of the world as understood by us. The basic concepts of metaphysics are the most basic structures of our understanding: any understanding of anything is structured by those concepts. So the world as understood by us is structured by those concepts, which means that those concepts refer: they refer to joints, not in the world as it is in itself, but in the world as it is understood by us.

The topic of transcendent metaphysics, by contrast, is the world as it is in itself, the structure that it has in its own right, not in relation to anything else. We have seen that both the pre-Kantian rationalists and some contemporary philosophers take this to be the appropriate topic of metaphysics. The point of Kant’s question—why do the basic metaphysical concepts refer?—is that the transcendent metaphysician cannot explain why her metaphysical concepts refer. Since the aim of metaphysics, on this view, is to correctly map the structure of the world in itself, and since that requires having concepts or terms that refer to that metaphysical structure, the transcendent metaphysician, if Kant is correct, cannot explain why transcendent metaphysics is possible. The practitioner of immanent metaphysics, on the other hand, can: since the structure of our understanding is the structure of the world as understood by us, then, if some set of concepts provides the structure of any possible understanding for any topic, those concepts refer to the joints in the world as understandable by us. To some, this may seem like it simply changes the topic: ‘immanent’ metaphysics is not really metaphysics at all. But if metaphysics is identified with transcendent metaphysics, then Kant’s conclusion is simply that the possibility of reference in metaphysics is inexplicable. Only in immanent metaphysics, the heir to traditional metaphysics, is the reference of our concepts explicable.

Kant raised a question—how is metaphysics possible?—that has multiple dimensions: he inquires into the epistemic, the scientific, the metaphysical, and the semantic possibility of metaphysics. I have focused on the semantic aspect of this question, which I take to mean: why do the basic concepts of metaphysics ‘relate to their objects’—that is, refer to them? I have argued that no one in the pre-Kantian or the post-Kantian tradition has a particularly successful answer to this question. Kant’s proposal is to reconceive metaphysics as the immanent science

of the structure of the framework in which we understand the world, and I have argued that we can explain how the basic concepts of that science refer. Kantian immanent metaphysics can answer the question about its own possibility that traditional transcendent metaphysics could not. If this does not constitute a great achievement in philosophy, I don’t know what does.

## Notes

- 1 See Förster (2012) for a discussion of what Kant meant by this astonishing claim. ‘Ak’ refers to the *Akademie* edition of Kant’s writings (Kant 1902), which I cite by volume and page number. The *CPR* I cite in the standard A/B format. Translations of the *CPR* are from Guyer and Wood (Kant 1998).
- 2 I borrow the term ‘cognitive semantics’ as a description of Kant’s theory from Hanna (2001).
- 3 Translation, with slight modifications, from Kant (1999: 133).
- 4 Kant alludes to the relation of empirical intuitions to their objects in the Herz letter itself (Ak. 10:130–1). He had outlined his theory of *a priori* intuitions and their relation to their objects (space, time) two years before in his ‘Inaugural dissertation’, *De mundi sensibilis atque intelligibilis forma et principiis* (Ak. 2:385–419; translation in Kant 1992).
- 5 I use angle brackets to denote the concepts that would normally be expressed by the italicized expression within: for example, <substance> is the concept of substance.
- 6 This is not the relation between <substance> and its extension, the set of substances. <Substance> would still have ‘its object’, even if it had a different extension (if one fewer or one more substance were to exist), so the ‘object’ of a concept (for Kant) is not its extension. I argue for this at length in Stang (forthcoming b).
- 7 Much of those debates were concerned with the reference of singular terms (names, definite descriptions) and indexical expressions; Putnam (1975) expands these issues to the case of natural-kind terms like ‘water’.
- 8 In the course of criticizing the ontological argument for the existence of God, Kant points out that there can be true analytic judgments involving concepts that are not instantiated by any objects (e.g., ‘God is omnipotent’ is true, even if there is no God). See A595–6/B622–3.
- 9 Obviously, there are complicated issues here about tense and context-sensitive expressions like indexicals and demonstratives, which, for reasons of space, I cannot discuss here.
- 10 This also allows us to extend the notion of reference beyond the reference of singular terms to objects and predicates to properties. For instance, if we adopt a primitivist modal metaphysics on which modal facts do not obtain in virtue of anything non-modal we can give the truth conditions of modal sentences as follows: ‘ $\diamond p$ ’ is true if and only if possibly  $p$ . In this case we could say that the modal operator  $\diamond$  ‘refers’ to possibility, since possibility facts contribute to truth conditions of sentences that involve it. Henceforth, to indicate that ‘referents’ can in principle be a broader category than objects I will speak of reference to ‘entities and structures’.
- 11 See Devitt and Sterelny (1999) and Davidson (2007) for an overview of the issues and the present state of the debate.
- 12 For instance, by Crusius’s theory of the highest material principles of thought; for discussion, see Heimsoeth (1956) and Hogan (2013).

- 13 Which of Hume's writings Kant had read, and, in particular, whether he had read the *Treatise*, is a matter of controversy among scholars. See Kuehn (1983); Kreimendahl (1990); and Beiser (2002: 43–7).
- 14 These are what Garrett (1997) calls 'revival sets' for ideas used abstractly.
- 15 Beiser (2002) offers this as further evidence that Hume had not read the *Treatise*.
- 16 In Stang (forthcoming *b*) I argue that Kant's locution '*the* object of a concept' never refers to the extension of a concept, much less to an individual object in that extension, but instead to what we might now call the 'content' of a concept.
- 17 See Locke (1975: bk. II, ch. XXI, §§1–4).
- 18 For example, Sider (2011: 162) and Williamson (2013: 423). Neither Sider nor Williamson would be attracted, though, to a Lockean 'empirical' deduction of the key terms in their metaphysical theories.
- 19 See Stang (forthcoming *a*) for extensive discussion of the modal content of Kantian experience (*Erfahrung*).
- 20 More precisely, the primary qualities of the body ground the object's power to produce in me ideas that resemble these qualities. In actual sense perception, these powers are activated and I come to perceive ideas that resemble qualities in bodies. Cf. Locke's *Essay* (1975: bk. II, ch. 8).
- 21 Careful readers will notice that this involves a transition from a claim of the form  $\Diamond\exists Fx$  to a claim of the form  $\exists(Fx \ \& \ \text{Possible}(x))$ , where  $\text{Possible}(x)$  is a predicate of objects rather than an operator on propositions. While many would reject the validity of this inference, I argue in Stang (2016) that Leibniz, Wolff, and Baumgarten are committed to the validity of this transition where  $F$  is an 'essential concept' (it picks out an object by its essential properties).
- 22 For ease of exposition I am going to be a bit cavalier about the exact ontological status of the referent here.
- 23 I reconstruct these arguments in detail in Stang (2016: ch. 3).
- 24 For critical discussion, see Chignell (2009); Stang (2010, 2016); Abaci (2014); and Yong (2014).
- 25 Note that this is not the question, 'Why (for what reason) did God create me with this concept?'
- 26 The derivative concepts are what Kant calls 'predicables of pure reason': see A82/B108.
- 27 Fictionalism about a problematic class of entities might also be thought of as eliminating reference to those entities: see, for instance, Rosen (1990) for a fictionalist account that *eliminates* reference to (Lewisian) possible worlds.
- 28 In forthcoming work, I explore the relation of metaphysical deflationism to Kant's own 'transcendental method' in metaphysics and I raise some Kantian (and post-Kantian) problems for the deflationist.
- 29 Fourth, there is the extension of the Concept, the set of things to which it correctly applies, but I will pass over that here.
- 30 I am assuming that 'water' expresses the same Concept in Earth English and in Twin-Earth English but refers to different things. But little in my argument hangs on this assumption; it is merely an illustration.
- 31 One could consider a view on which there are 'externalist' effects in the reference of metaphysical concepts (e.g., 'substance' refers to different things, depending on what world it is tokened in), but I will forgo discussion of that here.
- 32 This means that Kant is not committed to the implausible skeptical claim that, without a satisfactory explanation of the reference of, for instance, the terms of physics, our

- putative knowledge of physics is jeopardized. We might have knowledge in physics even if we lack second-order knowledge of how this knowledge is possible (among other things, why physical terms refer in the first place).
- 33 Nor do I think that these considerations are foreign to Kant. See his account of why we must assume that the 'system' of our empirical concepts must track the 'system' of nature in the Appendix to the 'Transcendental Dialectic' section of *CPR* A642/B670–A668/B696.
- 34 At this point I anticipate yet another buck-passing move in the assertion that explaining the reference of metaphysical terms is the job of semantics. But a semantic theory that explains the reference of terms in metaphysics is going to have to include quite a lot of metaphysics; whether this is called 'semantics' or 'metaphysical semantics' or the 'semantics of metaphysics' is a matter of indifference to me.
- 35 For the original idea of naturalness as an explanation of reference see Lewis (1984) in response to Putnam (1977, 1980, 1981). Sider (2011) is the most developed contemporary version of the idea that the world has a privileged metaphysical structure.
- 36 I owe this example to an unpublished paper by Laura Franklin-Hall.
- 37 There is now a small—but growing!—literature on *Erkenntnis* and how (and whether) it differs from knowledge, *Wissen*. See Smit (2000); Hanna (2006); Watkins and Willaschek (2017); and Schafer (forthcoming).

## References

- Abaci, Uygur 2014. Kant's Only Possible Argument and Chignell's Real Harmony, *Kantian Review* 19: 1–25.
- Beiser, Frederick C. 2002. *German Idealism: The Struggle against Subjectivism, 1781–1801*, Cambridge, MA: Harvard University Press.
- Chalmers, David, Manley, David, and Wasserman, Ryan (eds.) 2009. *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford: Clarendon Press.
- Chignell, Andrew 2009. Kant, Modality, and the Most Real Being, *Archiv für Geschichte der Philosophie* 91: 157–92.
- Davidson, Matthew (ed.) 2007. *On Sense and Direct Reference: Readings in the Philosophy of Language*, Boston: McGraw-Hill.
- Devitt, Michael and Sterelny, Kim 1999. *Language and Reality: An Introduction to the Philosophy of Language*, Cambridge, MA: MIT Press.
- Förster, Eckart 2012. *The Twenty-Five Years of Philosophy: A Systematic Reconstruction*, trans. Brady Bowman, Cambridge, MA: Harvard University Press.
- Garrett, Don 1997. *Cognition and Commitment in Hume's Philosophy*, New York: Oxford University Press.
- Hanna, Robert 2001. *Kant and the Foundations of Analytic Philosophy*, Oxford: Clarendon Press.
- Hanna, Robert 2006. *Kant, Science, and Human Nature*, Oxford: Clarendon Press.
- Heimsoeth, Heinz 1956. *Metaphysik und Kritik bei Chr. A. Crusius*, *Kant-Studien*, Ergänzungsheft 71, Köln: Kölner Universitäts-Verlag.
- Hofweber, Thomas 2009. Ambitious, yet Modest, Metaphysics, in David Chalmers, David Manley, and Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Clarendon Press: 260–89.
- Hogan, Desmond 2013. Metaphysical Motives of Kant's Analytic–Synthetic Distinction. *Journal of the History of Philosophy* 51: 267–307.

- Kant, Immanuel 1902–. *Kants Gesammelte Schriften*, 29 vols., Berlin–Brandenburg (formerly: Royal Prussian) Academy of Sciences, ed., Berlin: Walter de Gruyter.
- Kant, Immanuel 1992. *Theoretical Philosophy, 1755–1770*, ed. David Walford and Ralf Meerbote, Cambridge: Cambridge University Press.
- Kant, Immanuel 1998. *Critique of Pure Reason*, trans. and ed. Paul Guyer and Allen W. Wood, Cambridge: Cambridge University Press.
- Kant, Immanuel 1999. *Correspondence*, trans. and ed. Arnulf Zweig, Cambridge: Cambridge University Press.
- Kreimendahl, Lothar 1990. *Kant—der Durchbruch von 1769*, Köln: Dinter.
- Kuehn, Manfred 1983. Kant's Conception of 'Hume's Problem', *Journal of the History of Philosophy* 21: 175–93.
- Lewis, David 1984. Putnam's Paradox, *Australasian Journal of Philosophy* 62: 221–36.
- Locke, John. 1975 [1690], *An Essay Concerning Human Understanding*, ed. Paul Nidditch, Oxford: Clarendon Press.
- Putnam, Hilary 1975. The Meaning of 'Meaning', *Minnesota Studies in the Philosophy of Science* 7: 131–93.
- Putnam, Hilary 1977. Realism and Reason, *Proceedings and Addresses of the American Philosophical Association* 50: 483–98.
- Putnam, Hilary 1980. Models and Reality, *Journal of Symbolic Logic*, 45: 464–82.
- Putnam, Hilary 1981. *Reason, Truth and History*, Cambridge: Cambridge University Press.
- Rosen, Gideon 1990. Modal Fictionalism, *Mind*, 99: 327–54.
- Schafer, Karl forthcoming. Kant's Conception of Cognition and Our Knowledge of Things in Themselves, in *The Sensible and Intelligible Worlds: New Essays on Kant's Metaphysics and Epistemology*, ed. N.F. Stang and K. Schafer, Oxford: Oxford University Press.
- Sider, Theodore 2011. *Writing the Book of the World*, Oxford: Clarendon Press.
- Smit, Houston 2000. Kant on Marks and the Immediacy of Intuition, *Philosophical Review* 109: 235–66.
- Stang, Nicholas F. 2010. Kant's Possibility Proof, *History of Philosophy Quarterly* 27: 275–99.
- Stang, Nicholas F. 2016. *Kant's Modal Metaphysics*, Oxford: Oxford University Press.
- Stang, Nicholas F. forthcoming a. Hermann Cohen and Kant's Concept of Experience, in *Philosophie und Wissenschaft bei Hermann Cohen*, ed. Christian Damböck, Dordrecht: Springer.
- Stang, Nicholas F. forthcoming b. Kant and the Concept of an Object.
- Thomasson, Amie L. 2015. *Ontology Made Easy*, New York: Oxford University Press.
- Watkins, Eric and Willaschek, Marcus 2017. Kant's Account of Cognition, *Journal of the History of Philosophy* 55: 83–112.
- Williamson, Timothy 2013. *Modal Logic as Metaphysics*, Oxford: Oxford University Press.
- Yong, Peter 2014. God, Totality and Possibility in Kant's Only Possible Argument, *Kantian Review* 19: 27–51.