

# 4

## Carnap, the Necessary A Posteriori, and Metaphysical Anti-realism

*Stephen Biggs and Jessica Wilson*

### 4.1 Introduction

In *Meaning and Necessity* (1947/1950), Carnap advances an intensional semantic framework on which, as per typical empiricist assumption, modal claims are true in virtue of semantical rules alone, without reference to extralinguistic facts, and so are a priori. In ‘Empiricism, Semantics, and Ontology’ (1950), Carnap advances an epistemic-ontological framework on which, as per typical empiricist assumption, metaphysical claims are either trivial or meaningless, since lacking any means of substantive confirmation. Carnap carried out these projects two decades before Kripke influentially argued, in *Naming and Necessity* (1972/1980), that some modal claims are true, at least in part, in virtue of extralinguistic facts, and so are a posteriori. How should a neo-Carnapian respond to Kripke’s results? Some (notably, Chalmers and Jackson, in their joint 2001 and elsewhere) have suggested that an extension of intensional semantics along lines of “epistemic two-dimensionalism” can accommodate Kripke’s results while largely preserving commitment to the semantics-based a priority of modal claims. Here we consider how best to implement this suggestion, and how the resulting semantics fits with Carnap’s second project. We find that the most promising (and most Carnapian!) post-Kripke version of Carnap’s semantics—*abductive two-dimensionalism*—presupposes an epistemology which undermines Carnap’s metaphysical anti-realism.

### 4.2 The Project of *Meaning and Necessity*

Carnap aims, in *Meaning and Necessity*, to achieve two goals that are in the first instance empiricist but which may be seen as generally valuable: first, to provide an account of meaning that avoids certain metaphysical and semantic difficulties associated with Fregean “sense”; second, to use the associated semantic framework as a basis for interpreting and providing a logic for modal claims, in line with empiricist scruples.

Carnap aims to satisfy the first goal by constructing a semantic framework on which “to know the meaning of [a sentence] is to know in which of the possible cases it would be true and which not” (10); more generally, the suggestion is that the meaning of a given expression is given by its extension in each possible case or “state description”, where a state description is a maximal collection of sentences representing a (“Leibnizian”) possible world. Since expressions have extensions in *possible* cases, meaning involves more than actual extension, and accounting for meaning invokes *modality*. Carnap calls the non-extensional aspect of meaning ‘intension’.

Since we do not have experience of non-actual possibilities, how are we to identify the extensions of expressions in such cases, in order to identify their intensions? As we’ll discuss in Section 4.3.1, Carnap’s preferred strategy for associating intensions with expressions involves a pragmatically interpreted application of broadly abductive principles.<sup>1</sup> Independent of this strategy, there is a clear sense in which Carnapian intensions are an improvement over Fregean senses, from an empiricist point of view, in that access to an intension is ultimately a matter of access to (a range of) *extensions*, rather than a matter of rational or other grasp of a mind-independent abstractum.<sup>2</sup> Supposing that there is no in-principle problem for empiricists’ identifying the extension of a given expression given how the world actually is, one might naturally think that there is no in-principle problem for empiricists’ identifying the extension of a given expression given how the world might possibly be. The type of information is the same, after all: no new metaphysical category, requiring a potentially new form of epistemological access, is required.

Satisfaction of Carnap’s second goal—of providing an account of modal claims on which their truth is a matter of meaning rather than irreducibly modal mind-independent reality—emerges from satisfaction of the first. Intensions in hand, Carnap introduces *L-truth* as a specification of “what Leibniz called necessary truth and Kant analytic truth” (1947/1950, 8). The connection emerges as follows. A sentence is L-true just in case it is true in all state descriptions; furthermore, given that the intensions encoding what is true in state descriptions do so purely as a matter of meaning, it follows that a sentence is L-true just in case it follows from “semantical rules . . . alone without any reference to (extra-linguistic) facts” (1947/1950, 10). Continuing on: a sentence is L-false just in case it is false in every state description. One sentence L-implies another just in case the latter is true in every state description in which the former is true. Two sentences are L-equivalent just in case the sentence expressing their equivalence is L-true. And a sentence is L-determinate just in case it is either L-true or L-false, where

<sup>1</sup> Reflecting the pragmatic supposition, when engaging in exegesis of Carnap’s view we speak of speakers’ “identifying” rather than of their “knowing” which intension is associated with a given expression.

<sup>2</sup> Hence notwithstanding that Carnap says that “Frege’s concept of sense is very similar to that of intension” (129), he also cites the usual empiricist concerns with reification of abstracta, and highlights that the concepts he appeals to “do not involve a hypostatization” of the sort that Frege associated with sense, since “our statements belong to, or can be translated into, the general language of science” (22)—that is, can be cashed in terms of experienced extensions.

the contrast here is with “L-indeterminate” claims that are “factual”, “synthetic”, or “contingent”.<sup>3</sup> Carnap maintains that every modal sentence is L-determinate; hence on his view the truth of every modal sentence follows from semantic rules alone, independent of extralinguistic facts, as empiricists typically suppose.

### 4.3 The Necessary a Posteriori and Epistemic Two-Dimensionalism (E2D)

Kripke’s (1972/1980) insights threaten to undermine Carnap’s account of meaning and modality. Kripke rejects both descriptivist theories of meaning (of the sort tacitly presupposed in Carnap’s intensional semantics) and the conflation of necessity with a priority. He argues, more specifically, that some names and natural kind predicates do not correspond to cognitively accessible reference-fixing descriptions, but are rather to some extent directly referential, such that certain modal claims involving such expressions can be known only a posteriori. If Kripke is right, as we suppose in what follows, then it seems Carnap must be wrong.

Despite this threat, one might think that there is no deep difficulty for post-Kripke Carnap here, for as Jackson (1998), Chalmers and Jackson (2001), and Chalmers (2006) (among others) have argued, Kripke’s insights can be preserved within a broadly descriptivist, intensional semantic framework, consonant in large part with the supposition that what is necessary is a matter of meaning, by means of *epistemic two-dimensionalism* (E2D).

The basic E2D strategy for neo-Carnapian accommodation of a posteriori necessities is as follows. To start, the suggestion is that even though we cannot know all modal claims a priori, we (or idealized versions of us) can have fairly comprehensive, semantically-based, a priori knowledge of the intensions underlying all necessary truths, including necessary a posteriori truths.<sup>4</sup> In particular, for any necessary a posteriori truth  $T$ , knowing the semantic rules governing a sentence that expresses  $T$ , which includes knowing the rules for each sub-sentential expression comprising that sentence, puts one in position to know two conditionals that together serve as the ultimate foundation of  $T$ ’s truth.

So, for example, using only our knowledge of semantic rules, we can discover that (i) if we are in a world that would make true a state description according to which the watery stuff is  $H_2O$ , then the actual extension of ‘water’ is  $H_2O$ , and (ii) if the actual extension of ‘water’ is  $H_2O$ , then  $H_2O$  is the extension of ‘water’ in all possible worlds. Likewise, *mutatis mutandis*, for worlds that would make true state

<sup>3</sup> “In general, factual knowledge is needed for establishing the truth-value of a given sentence. However, if the sentence is L-determinate [ ... ], the semantical rules suffice for establishing its truth value or, in other words, its extension” (69).

<sup>4</sup> Note that, as per the “epistemic” in “epistemic two-dimensionalism”, the strategy departs from Carnap’s supposition that the association of intensions with expressions is a pragmatic matter. We follow up on this difference below.

descriptions according to which the watery stuff is other than  $H_2O$ . According to E2D, then, all knowledge of modality is ultimately grounded in meaning, with semantically based a priori knowledge of conditionals providing the bridge from empirical contingencies (e.g., that the actual watery stuff is  $H_2O$ ) to a posteriori necessities (e.g., that water is necessarily  $H_2O$ ). On this picture, the only role for empirical investigation in modal knowledge lies in determining which world is (or which relevant non-modal facts are) actual; this fact (these facts) in hand, we can then discharge the antecedent of the relevant a priori conditionals, and gain access to the unconditional necessary truths.

The semantics at issue in E2D is *two-dimensional* in that many expressions, including those that figure in a posteriori necessities, have two interrelated intensions, associated with the two types of conditionals above. The *primary* intension of an expression  $E$  is a function that takes as input any state-description  $s$  (now representing a centered possible world, or scenario<sup>5</sup>), and delivers as output the extension  $E$  would have if  $s$  were actually true—hence, (i). The *secondary* intension for an expression  $E$  is a function that takes as input both the state-description  $s$  that is actually true (again, representing a centered world, or scenario), and any world  $w$ , and delivers as output the extension of  $E$  at  $w$ . We find it useful to introduce another intension: the *generalized secondary intension* for an expression  $E$  takes as input any (centered) state description  $s$  and any world  $w$ , and delivers as output the extension  $E$  would have at  $w$  if  $s$  were actual—hence, (ii). The E2D strategy presupposes that both primary intensions and generalized secondary intensions can be known a priori, on the basis of semantic competence alone. Hence, on this view, our semantically based knowledge of the primary and generalized secondary intensions for natural kind expressions provides an ultimate foundation for our knowledge of ordinary secondary intensions and corresponding a posteriori necessities, even though only experience can discharge the antecedent of conditionals such as (i).

This much a posteriority is arguably compatible with the empiricist tenet that modal claims are true in virtue of meaning—at least, we suppose here that this is correct. One could think of this supposition as regulating what we mean by ‘grounds’ when we say that semantically based knowledge *grounds* all modal knowledge. More generally, this much a posteriority is arguably compatible with our having, as is desirable independently of empiricism, significant access to the space of possibility, prior to the end of empirical inquiry.

#### 4.3.1 Carnap’s intensional semantics and E2D

Does E2D in fact provide a suitably neo-Carnapian accommodation of a posteriori necessities? In order to answer this question, we need first to acknowledge and assess two potential “mismatches” between Carnap’s semantics and E2D.

<sup>5</sup> A centered world, or scenario, is effectively a world along with an indexical perspective, allowing for meaning to reflect, e.g., the facts “around here”.

The first pertains to Carnap's supposition that the association of intensions with expressions is a pragmatic rather than epistemic matter, in contrast with the epistemic interpretation associated with any version of E2D, which aims not just to conform to empiricism, but also to make generally desirable sense of our epistemic access to a suitably wide range of modal claims. Chalmers and others assume that taking the assignment of intensions to expressions to be an epistemic matter doesn't prevent the resulting semantics from being properly "Carnapian"; and in the text to follow we also take this for granted, in order to more directly consider the bearing of the E2D strategy on Carnap's metaphysical anti-realism. In Appendix A to this chapter, however, we consider Carnap's reasons for taking intensions to be pragmatically determined, and argue that they are unconvincing, and are moreover undermined by Kripke's results.

The second and more important potential mismatch, which will mainly concern us in what follows, is between Carnap's account of "explication" as the preferred methodology for identifying intensions, and the conceiving-based approach presupposed by Chalmers and Jackson (2001) in their more-or-less standard interpretation of E2D. In Section 4.4 we present these differing approaches to one's knowledge of intensions, highlighting the abductive nature of explication and Carnap's reasons for thinking that the distinctive features of abduction are needed to overcome the widespread "vagueness" of natural kind expressions/concepts; we then argue that an abductive approach is indeed required for this purpose. Correspondingly, we maintain that the most promising—and most Carnapian—implementation of the E2D strategy relies on *abduction* rather than *conceiving* as the preferred epistemology of intensions.

Perhaps the main concern with interpreting E2D in this fashion cites the supposition that the results of abduction are not a priori; in Section 4.5 we address this objection, arguing that the results of idealized abduction are reasonably seen as being a priori, and in particular, as being a priori as the results of idealized conceiving. The results of Sections 4.4 and 4.5 then provide setup for the discussion, in Section 4.6, of how a suitably neo-Carnapian accommodation of the necessary a posteriori bears upon Carnap's metaphysical anti-realism.

#### 4.4 Conceiving vs. Abduction as the Epistemology of Intensions

Chalmers and Jackson interpret E2D as involving a conceiving-based epistemology of intensions, according to which the association of intensions with expressions can and should proceed by means of conceiving.

What is conceiving? Chalmers and Jackson take conceivability and a priority to go hand in hand, such that a sentence token is conceivable if and only if it is not ruled out a priori. Chalmers elaborates: "a sentence token is a priori when it expresses an a priori thought", where an a priori thought is one that "can be conclusively non-experientially justified on idealized rational reflection" (2006, 98); a thought is justified conclusively

if its actual justification ensures the truth of any sentence expressing it; and a thought is justified non-experientially if it is justified on the basis of idealized rational reflection alone. The reflection at issue is (suitably) idealized if any belief that it produces corresponds to a sentence that can be known by a hypothetical thinker who can entertain any scenario, possesses exactly the concepts and language that we possess, can know whatever can be known through rational reflection on the same, and can know nothing else. (Henceforth, such idealization is presumed.) So, a sentence token is conceivable if and only if the thought expressing its denial is not (conclusively) justified on the basis of rational reflection alone, and is inconceivable otherwise.

What is “rational reflection”? The rough answer in Chalmers and Jackson (2001) is that rational reflection is a process in which implicit conceptual analysis manifests as explicit judgments about the extensions of one’s expressions at scenarios or scenario-world pairs.<sup>6</sup> How does this process work? Most important for our purposes is that, in contrast to the method that Carnap prefers, Chalmers and Jackson are explicit that rational reflection excludes appeals to theoretical virtues (2001, 342).<sup>7</sup> They are less clear about the positive details, though the intended contrast with theoretical virtues, along with the supposition that the results of conceiving are “conclusive”, suggests that the process involves some infallible analogue to perception or intuition, enabling one to deduce, see, or intuit the contents of and relations among concepts.

#### 4.4.1 *Indeterminacy and conceiving*

Why does Carnap reject a conceiving-based method of identifying intensions? Roughly, Carnap thinks both that many natural kind expressions are indeterminate,<sup>8</sup> and that conceiving cannot resolve this indeterminacy; these commitments in turn imply that conceiving cannot ground our access to a wide range of intensions and corresponding modal truths. Since we agree with Carnap, we will develop this position on his behalf. Specifically, we will argue that many natural kind expressions/concepts are indeterminate, and that attempts to overcome this indeterminacy by a conceiving-based epistemology of intensions fail; we will then explain why such failures render conceiving unsuited for purposes of implementing the E2D strategy, and more generally, unsuited for any intensional semantics aiming to ensure access to an appropriately wide range of modal truths.<sup>9</sup>

<sup>6</sup> We often cite Chalmers and Jackson’s joint paper (2001), which explores conceptual analysis and reductive explanation, not two-dimensional semantics, because it clarifies issues related to E2D.

<sup>7</sup> See Biggs and Wilson (forthcoming) for discussion of and objections to Chalmers’ and Jackson’s reasons for thinking this.

<sup>8</sup> Carnap rather uses the term “vague”, but we will follow common practice of using “indeterminacy” to refer to the general phenomenon, restricting “vagueness” to cases of indeterminacy involving sorites-susceptible expressions/concepts. The use of ‘indeterminacy’ here has nothing to do, of course, with Carnap’s talk of ‘L-indeterminate’ (broadly contingent) statements.

<sup>9</sup> Following Carnap, we move freely from discussing expressions, which we take to have their standard interpretations, to discussing the concepts that they express; hence our use of ‘expressions/concepts’.

We start with Mark Wilson's (1982, 2006) claim that many natural kind predicates are indeterminate.<sup>10</sup> The general idea behind Wilson's claim is that such indeterminacy is indicated by a seeming arbitrariness of application of natural kind predicates to new cases. To take his illustrative example, members of an isolated tribe might or might not include airplanes in the extension of 'bird', depending on whether they happen to first encounter an airplane overhead or on the ground. If whether 'bird' applies to the airplane depends on historical accident, then, Wilson plausibly claims, the full range of extensions of the expression is not antecedently determined.

The historical record supplies other cases where application of an expression reflects factors whose influence is not plausibly seen as antecedently encoded in the expression/concept. 'Acid' initially was taken to refer to only oxygenated substances, but was later applied to HCl, for theoretical reasons now largely discarded; dispute remains over whether Newtonian uses of 'mass' apply in relativistic contexts; the decision to classify whales as mammals was a controversial affair; and there was a recent resolution declassifying Pluto as a planet. Ordering phenomena in Sorites series also suggest that arbitrary or extrinsic factors can influence predicate application; the breaking point in applications of 'blue', for example, may depend non-systematically both on where in the spectrum one starts and on psychological factors (cf., Raffman 1994). Summing up: sometimes the factors influencing predicate application will involve historical accident, as in Wilson's toy case; sometimes they will involve non-demonstrative reasoning, as for 'acid', 'mass', 'mammals', and 'planet'; sometimes they will involve variable psychological features. In all these sorts of situation, it appears that decisions to apply (or not) the predicate at issue depend on factors whose influence is not antecedently encoded in that predicate.<sup>11</sup>

To see the challenge that such widespread indeterminacy poses to E2D, suppose momentarily that *every* natural kind expression is partly, insuperably indeterminate. In that case, for any natural kind expression *E*, there is a state description at which we cannot identify the extension of *E*. As such, we can identify neither the primary intension nor the generalized secondary intension of *E*—since these take us to the extension of *E* at *each* state description, and state-description/world pair. But according to E2D, knowledge of these intensions grounds knowledge of modal truths, including truths involving natural kinds. Hence if every natural kind expression is partly, insuperably indeterminate, E2D cannot explain our access to modality.

The point can be made another way. The "core thesis" of E2D (Chalmers 2006a, §3.1) is that a sentence is a priori just in case its primary intension is true at every scenario;

<sup>10</sup> Here we draw upon our discussion in (Biggs and Wilson, in progress *a*).

<sup>11</sup> Distinguishing expressions and concepts for the moment: one might see such cases as suggesting that, while natural kind predicates determinately express a certain concept, many such concepts are indeterminate, in that they may not, in and of themselves, determine their application (or that of associated predicates) to the full range of scenarios; alternatively, one might see such cases as suggesting that it is indeterminate which of multiple completely determinate concepts a given natural kind predicate expresses. Either way, the difficulty to be next discussed will apply.

this thesis in turn implies that a sentence is not a priori if its primary intension is insuperably indeterminate at even one scenario. It follows that if all natural kind expressions are partly, insuperably indeterminate, then no sentences are a priori, and thus, contrary to E2D, our knowledge of modal claims involving natural kinds cannot be grounded in semantically based, a priori knowledge of intensions.

To be sure, the claim that *all* natural kind terms are indeterminate is likely too strong. Even supposing some such terms are determinate, however, the moral of the previous discussion is that on E2D, the more insuperable indeterminacy there is, the fewer modal truths we are in position to know. As such, E2D's primary goal—of making sense of our having (in principle) knowledge of a *wide range* of modal truths, including those involving natural kinds, is incompatible with widespread, insuperable indeterminacy of natural kind expressions. Any significant degree of indeterminacy, then, poses a serious challenge to E2D, and to any intensional semantics with similar aims.

Chalmers claims, nonetheless, that such indeterminacy presents “no problem” for E2D:

There may of course be borderline cases in which it is indeterminate whether a concept would refer to a certain object if a given world turned out to be actual. This is no problem: we can allow indeterminacies in a primary intension, as we sometimes allow indeterminacies in reference in our own world. (1996, 364)

Chalmers is right that E2D can tolerate some indeterminacy, such that the primary intension of some expressions cannot be known a priori. But as above, E2D cannot *both* allow that indeterminacy is widespread, *and* explain our access to a wide range of modal truths.

Anticipating this difficulty, Chalmers suggests that conceivers can eliminate indeterminacy from primary intensions by foreseeing relevant accidents. For example, in re Wilson's toy case, Chalmers says that conceivers “might try to classify these two different scenarios [airplane first seen in the sky or on the ground, respectively] as different ways for the actual world to turn out, and therefore retain a fixed, detailed primary intension” (1996, 364). On this broadly supervenience strategy, the fully determinate primary intension of ‘bird’ includes planes in its extension if the tribe members first see a plane overhead but not if they first see it grounded. Either way, according to Chalmers, the indeterminacy is resolved.

Chalmers' suggestion has potential re Wilson's concerns only if a conceiver can foresee how intensions are sensitive to accidents. But as we see it, a deeper lesson of Wilson's case is that the influence of accidents cannot be foreseen. Determinism and such aside, there might be divergence of application even relative to the same historical facts; after all, there are any number of respects of dissimilarity between airplanes and birds, even when the former are in flight, and a minor difference in attention to these features (or even mood) might result in a different decision about whether ‘bird’ applies. We can register, post hoc, extensions resulting from whatever decision was in fact made; but

why think that idealized conceivers would be in position to antecedently identify the corresponding extensions and intensions? Moreover, Chalmers' suggestion only addresses cases where historical accident influences decisions about extension. As above, other factors may similarly undermine the supposition that intensions are antecedently encoded in concepts, as when, for example, theoretical virtues enter into decisions about how to classify HCl, Newtonian mass, whales, and Pluto.<sup>12</sup>

In presupposing a conceiving-based epistemology, then, Chalmers and Jackson's version of E2D fails to be appropriately Carnapian, not just in ignoring Carnap's preferred method for identifying intensions, but also in failing to address the legitimate concerns about conceptual indeterminacy leading Carnap to that method—concerns that have only gained in support since Carnap's time. E2D understood as involving a conceiving-based epistemology cannot accommodate the necessary a posteriori, and thus fails to achieve its primary aim.<sup>13</sup>

#### 4.4.2 Carnap's abductive route to intensions

Does an intensional semantics that presupposes Carnap's preferred method fare better? We begin to answer this question by sketching his method, as found in his account of explication. Roughly, explication is "making more exact" (which Carnap understands as "replacing") a "vague or not quite exact concept" with a "newly constructed, more exact concept" (1947/1956, 7–8). Although explication is central to Carnap's semantics, *Meaning and Necessity* offers few details about the method, about how one makes a concept more exact or chooses an appropriate replacement. Instead, explication is there introduced through illustrative examples, as when Carnap offers L-truth as the result of explicating "logical or necessary or analytic truth" (1947/1950, 7).

In *Logical Foundations of Probability* (1950), Carnap provides the needed details, opening with a chapter on explication. He first reiterates what explication is:

The task of *explication* consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second. We call the given concept (or the term used for it) the *explicandum*, and the exact concept proposed to take the place of the first (or the term proposed for it) the *explicatum*. (italics in text, 1950, 3)<sup>14</sup>

<sup>12</sup> One should not take theoretical virtues to be built into intensions such that, e.g., if an appeal to fruitfulness pushes chemists to apply 'acid' to HCl at some scenario, then it follows that the intension of 'acid' includes that virtue, and gives it special importance. For building theoretical virtues into intensions radically multiplies associated concepts, requiring a distinct concept for each combination of virtues. Moreover, this implausible result has the implausible consequence that most if not all disagreement about the extension of one's expressions is non-substantive; see Biggs and Wilson in progress *a* for further discussion.

<sup>13</sup> These concerns about indeterminacy are not the only challenges for E2D. In Biggs and Wilson in progress *a*, we argue that a range of seemingly compelling objections to E2D, including those due to Byrne and Pryor (2006), Schroeter (e.g., 2004), and Block and Stalnaker (1999), only target E2D when implemented using a conceiving-based epistemology of intensions, and that given an abduction-based epistemology of intensions of the sort offered below, E2D can meet such challenges.

<sup>14</sup> Carnap cites Kant and Husserl as inspirations for his use of 'explication' (1950, 3): for Kant, judgments that affirm analyticities are 'explicative'; for Husserl, 'Explicat' are precisifications of confused, unarticulated senses. Carnap's take resembles Husserl's more closely than Kant's in that Kant thinks of explications

Carnap then turns to how explication works or, equivalently, to what makes one explicatum for a given explicandum superior to another. He begins with four conditions that a “concept must fulfil . . . in order to be an adequate explicatum for a given explicandum: (1) similarity to the explicandum, (2) exactness, (3) fruitfulness, (4) simplicity” (1950, 5). Although Carnap offers (1)-(4) as mere conditions on adequacy, he treats them as criterial, so that for any explicandum  $D$  and any explicata  $T$  and  $T^*$ , if  $T$  is most similar to  $D$ , most exact, most fruitful, and most simple, then one should choose  $T$  over  $T^*$  as the explicatum for  $D$ . Taken together, then, these conditions constitute a method for choosing among competing intensions for a given concept. Of course, an explicatum may be superior to alternatives in one respect but inferior in another— $T$  may be most similar to  $D$  while  $T^*$  is most fruitful, for example. Accordingly, one needs a way to balance the competing criteria. Carnap provides a rough account, according to which fruitfulness is paramount.

On Carnap’s account of explication, then, one chooses among competing intensions for a given expression (i.e., competing explicata for a given explicandum) by using *theoretical virtues*, balanced in a particular way. We take this method to be an instance of inference to the best explanation—that is, of *abduction*. Abduction, as we think of it, proceeds by assessing the extent to which a range of candidate theories satisfies the (perhaps competing) dictates of various theoretical virtues—parsimony, comprehensiveness, fruitfulness, and so on. To use abduction when deciding among competing theories is to choose the theory (*explanans*, *explicatum*) that best explains some target (*explanandum*, *explicandum*), where underlying theoretical virtues, appropriately balanced, determine how theories are ranked. Accordingly, Carnap offers an abduction-based method for identifying intensions.

How can abduction, so characterized, help one choose among competing intensions? Answering this question requires identifying the theories and targets at issue, and then showing how appealing to theoretical virtues can help one choose among those theories.

We take theories of intensions to be the candidate intensions themselves, i.e., the competing explicata. Candidate primary intensions for ‘water’, for example, might hold that, in scenarios considered as actual, ‘water’ refers to, respectively, (i) the basis of life; (ii) the watery stuff; (iii)  $H_2O$ . These theories might aim to explain, among other things, what we would take the extension of ‘water’ to be if the actual world had turned out to be one where the watery stuff was perfectly coincident with XYZ rather than  $H_2O$ . How might theoretical virtues enter into ranking these theories of the primary intension of ‘water’? Most saliently, the theory in (iii) is in one respect *less explanatorily*

as merely decomposing explicanda, as identifying the predicates already contained therein, while Husserl thinks of explications as (potentially) extending beyond explicanda, albeit in a principled way. Carnap also suggests that his use of ‘explication’, ‘explicandum’, and ‘explicata’ resemble Langford’s use of ‘analysis’, ‘analysandum’, and ‘analysans’; that his views on explication resemble Moore’s views on analysis, as articulated by Schillip; and that his thinking about explication resembles Naess’ thinking about ‘precisation’ (8). Beaney (2004) suggests that Frege’s views on analysis also may have (perhaps indirectly) influenced Carnap.

*comprehensive* than its competitors, since it cannot explain the thoughts of those contemplating hypothetical scenarios, or the actions of those in hypothetical scenarios, in which the watery stuff is/is coincident with any substance other than H<sub>2</sub>O.

Similarly, we take theories of generalized secondary intensions to be the candidate intensions themselves. The candidate secondary intensions of ‘water’, for example, might express that, in any world considered as counterfactual relative to an “H<sub>2</sub>O-scenario” considered as actual, ‘water’ refers to, respectively, (i) the basis of life; (ii) the watery stuff; (iii) H<sub>2</sub>O. These theories might aim to explain, among other things, that the actual extensions of ‘water’ and ‘H<sub>2</sub>O’ perfectly coincide. How might theoretical virtues enter into ranking these theories? Plausibly, the theory at issue in (iii), in identifying the secondary intensions of ‘water’ and ‘H<sub>2</sub>O’, explains the perfect coincidence of the actual extensions of these expressions in a more *ontologically parsimonious* way than its competitors.<sup>15</sup>

An abduction-based method for identifying intensions, then, provides a basis for choosing among competing intensions, no less than a conceiving-based method.<sup>16</sup>

#### 4.4.3 *The widespread indeterminacy of natural kind expressions, and the need for abduction*

An abduction-based method, moreover, succeeds where a conceiving-based method fails, in overcoming conceptual indeterminacy. Consider Wilson’s toy case. When deciding how to apply an expression in a given scenario, abductors can consider not only historical accident and psychological variability, but also any non-demonstrative rational considerations that might push one way or another, for theoretical virtues can encode any such considerations. Accordingly, abduction, unlike conceiving, is potentially *productive*. Consequently, abductors need not rely, post hoc, on historical or other facts along the way to identifying intensions, but may consider, even independent of such facts, what decisions *would* or *should* be made, through the proper use of abduction. More broadly, since abduction can rationally transcend what expressions antecedently encode, an abduction-based method has the potential to overcome each of the varieties of indeterminacy discussed earlier, extending applications of natural

<sup>15</sup> In developing an abduction-based modal epistemology, Biggs (2011) considers how claims about necessity and contingency provide (better or worse) explanations of various facts. That work transfers readily to the present discussion. Importantly, the above sketch leaves open which virtues are at issue, and how they should be balanced. We leave these details open both because one can see how abduction can resolve indeterminacy without entering into such details, and because no specific set of virtues or way of balancing is uncontroversial. Such flexibility, in our view, is a feature, not a bug, of this method. For more on abduction, see Lipton (1991/2004).

<sup>16</sup> Lavers (this volume) also offers a detailed discussion of explication, in the course of arguing that Carnap’s support for ontological relativism in “Empiricism, Semantics, and Ontology” turns on the claim that neither ‘truth’ nor ‘reference’ has a unique explication. Kraut (this volume) also discusses explication as “a meaning analysis or a reductive account of truth conditions”, intimating that it is, “content-preserving” (p. 37). As we note in Appendix A, Carnap initially suggests that explication needn’t be content-preserving, but in his reply to Strawson he is ambivalent about how content-preserving explication must be.

kind expressions to new scenarios, on ultimately rational grounds. While more could be said about this issue, we suspect that similar considerations drove Carnap's preference for an abduction-based method, and as such we anticipate that he would, and neo-Carnapians should, find these considerations compelling.

Let's sum up the results thus far. The initial question that concerns us is whether knowledge of necessary a posteriori truths can be accommodated within a broadly Carnapian framework on which modal claims are true in virtue of semantical rules, known a priori. The E2D strategy seems well-suited for this purpose, but the usual understanding of this strategy, as relying on a conceiving-based epistemology of intensions, is at odds with Carnap's explication-based means of identifying intensions, and in any case is moreover unable to overcome the conceptual indeterminacy that motivated Carnap's reliance on explication in the first place. If, however, the E2D strategy is implemented using an abduction-based epistemology of intensions, such indeterminacy can be overcome, in a way consonant with Carnap's explication-based approach. So far, so good, then, for a neo-Carnapian treatment of the necessary a posteriori.

## 4.5 The A Priori Status of the Products of Abductive Deliberation

Perhaps the most pressing objection to the suggestion that a properly Carnapian version of E2D should appeal to an abductive rather than a conceiving-based epistemology of intensions is that E2D requires that our access to intensions be a priori; but, it is claimed, the results of abductive deliberation are a posteriori. We reply, perhaps surprisingly, that the results of abduction are appropriately a priori.<sup>17</sup>

What makes knowledge *a priori*? To start, note that experience can play four different roles in knowledge formation. Say one knows that *p*. Then...

1. Experience might play a role in acquiring the concepts that are required to think *p*.
2. Experience might play a role in knowing the evidence that is required to know that *p*.
3. Experience might play a role in knowing that the inferential procedures deployed in coming to know that *p* are epistemically significant (i.e., have justificatory force).
4. Experience might play a role in acquiring/coming-to-use those procedures.

If knowing that *p* does not involve experience's playing any of the roles in 1–4, then that knowledge would be a priori. But can knowledge be a priori even if experience plays an inescapable role along some of 1–4?

<sup>17</sup> In (Biggs and Wilson forthcoming) we additionally consider an objection according to which Chalmers and Jackson's reason for excluding theoretical virtues from conceiving undermines an abduction-based epistemology of intensions; we reply that their assessment rests on a misunderstanding of what theoretical virtues are, and thus, does not threaten an abduction-based epistemology of intensions.

Let's first consider how advocates of a conceiving-based epistemology answer this question. In re 1: advocates of a conceiving-based epistemology follow the crowd in allowing that knowledge can be a priori even if experience is needed to acquire relevant concepts. For example, knowledge that bachelors are unmarried can be a priori even if experience must play a role in acquiring concepts expressed by 'bachelor' and 'male'. In re 2: advocates of a conceiving-based epistemology maintain that knowledge that  $p$  is a posteriori if experience must play a role in acquiring the evidence required to know  $p$ . For example, knowledge that water is necessarily  $H_2O$  is a posteriori, since experience plays an inescapable role in knowing that the watery stuff is actually  $H_2O$ . That said, advocates of a conceiving-based epistemology maintain, as per the discussion in S2, that such a posteriori knowledge is largely "grounded in the a priori": while one cannot know that 'water' refers to  $H_2O$  in all possible worlds without learning through experience that water is actually  $H_2O$ , nonetheless the conditional claim that *if* water is actually  $H_2O$ , then it is necessarily so, falls out of relevant intensions which are known a priori.

In re 3: advocates of a conceiving-based epistemology maintain that knowledge that  $p$  is a posteriori if experience plays a role in establishing that the procedures deployed in coming to know that  $p$  are epistemically significant, and they deny that experience is needed in order to establish the epistemic significance of conceiving. In re 4: advocates of a conceiving-based epistemology do not, so far as we can tell, explicitly address whether experience plays a role in learning how to conceive. That said, it would not be surprising if experience *does* play an important role in learning to conceive—as we all know, our students often need encouragement to think in an appropriately imaginative way about what is possible rather than about (just) what is actual. In any case, that advocates of a conceiving-based epistemology do not explicitly come down *against* experience playing this role itself suggests that even if it does, they would (reasonably, we think) not take this to undermine their claim that the products of conceiving are a priori.

How do advocates of an abductive epistemology of intensions—how do we—view the bearing of roles 1–4 on a priority? To start, we take exactly the same stance on 1 and 2 as do advocates of a conceiving-based epistemology. In re 1: we maintain that knowledge can be a priori even if experience is needed to acquire relevant concepts. In re 2: we maintain that knowledge is a posteriori if experience must play a role in acquiring the evidence required to know it, although some such knowledge—in particular, knowledge of a posteriori necessities—is nonetheless largely grounded in the a priori.

We also agree with advocates of a conceiving-based epistemology in re 3: we maintain that knowledge that  $p$  is a posteriori if experience plays a role in establishing that the procedures deployed in coming to know that  $p$  are epistemically significant. So, given our advocacy of an abductive epistemology of intensions, we must accept that the epistemic significance of abduction and underlying theoretical virtues can be established a priori.

While this claim is uncommon, we think there is good reason to accept it, on the broadly transcendental ground that the epistemic significance of abduction is a

necessary precondition for the possibility of right reasoning. The transcendental claim is motivated, in turn, by its being the case that the choice of a theory  $T^*$  over a competing theory  $T$  scoring at least as well, and in some cases better, on every theoretical virtue, would clearly be *irrational*. So, for example, the choice of a theory  $T^*$  over a competing theory  $T$ , where  $T^*$  and  $T$  score equally well on all theoretical virtues *except* that  $T^*$  is more convoluted than  $T$ , would clearly be irrational. Moreover, such a choice would be irrational *no matter what the world was like*. It is correspondingly impossible to think of right reasoning as proceeding via a principle that, other things being equal, one should choose the most convoluted theory; and similarly for other counter-abductive principles. Hence we can know a priori that “counter-abduction” isn’t epistemically significant. Similarly, the choice of a theory  $T$  that does as well as and sometimes better than its competitors, on every theoretical virtue, would clearly be *rational*—and such a choice would be rational *no matter what the world was like*. Hence we can know a priori that abduction and underlying theoretical virtues are epistemically significant.<sup>18</sup>

We think, though not everyone may follow us in this, that the above asymmetry is illuminated by the broadly transcendental supposition that both abduction and underlying theoretical virtues are *constitutive of human reasoning*—are as core to right thinking as principles of logical inference (c Kant 1781/1998)—so that the epistemic significance of abduction and underlying theoretical virtues is a necessary precondition for the possibility of human reasoning. The claim that abductive inference is constitutive of human reasoning is, moreover, independently plausible, in being supported by considerations from cognitive psychology (Gelman and Markham 1986; Feeney and Heit 2007).<sup>19</sup> We will address an objection to our claim that experience is not needed to establish the epistemic status of abduction shortly, but we take it that, antecedent to down-the-line objections, the previous considerations serve as reasonable *prima facie* motivation for our position on 3.

Finally, in re 4: we maintain, as advocates of a conceiving-based epistemology may do, that even if some experience is required in order to acquire or apply the inferential process at issue, this need not impugn the status as a priori of the deliverances of the process. To be sure, there is a difference with a conceiving-based epistemology here, since notwithstanding that we are natural born abductors (as per Gelman and Markham 1986; Feeney and Heit 2007), experience can surely *tweak parameters* associated with abductive principles, as it can affect the sample size one requires for inductive generalization. But, importantly, such tweaking is compatible with

<sup>18</sup> One might object that it would or could be rational to choose the most convoluted theory in any world where convoluted theories tend to be true. In Biggs and Wilson (in progress *b*), we address this objection, arguing that it rests on a mistake, in failing to distinguish facts from normative epistemic principles.

<sup>19</sup> These considerations also indicate that the sense of ‘transcendental’ at issue here is compatible with a naturalist worldview, according to which philosophical investigations are broadly continuous with those of the sciences. See also the analogy to the principles and parameters account of grammar, to be shortly discussed.

abduction's being an innate, and indeed necessary, component of our rational cognitive economy.

Compare the principles and parameters approach to grammar (cf. Chomsky and Lasnik 1993), according to which experience can tweak the parameters of an innate grammar. In the latter case, the role played by experience involves giving broadly contingent content to parameters in grammatical structures that are in the relevant sense necessary, in being determined independently of experience and common to all speakers of anything we would recognize as language. Similarly, in the case of abduction, the role played by experience involves giving broadly contingent content to parameters in epistemic structures that are in the relevant sense necessary, in being determined independently of experience, and common to all thinkers engaging in anything we would recognize as right reasoning. Closer to home, we see the role played by experience in tweaking abductive parameters as relevantly like that played in the acquiring of concepts: in both cases, experience fills in certain aspects of the *content* needed to engage in epistemic deliberation, without undermining the broadly *formal* or *structural* reasons (such as those we gave above for abduction and associated theoretical virtues, in re 3) for thinking that the products of such deliberation are a priori. We thus maintain that even if experience plays a role in re 4, this does not prevent the products of abduction from being a priori.<sup>20</sup>

Summing up: the deliverances of an abduction-based epistemology are reasonably taken to be a priori, and, moreover, are reasonably taken to be as a priori as the deliverances of a conceiving-based epistemology. If conceiving can deliver a priori knowledge of intentions, then so can abduction.<sup>21</sup>

<sup>20</sup> Williamson (2007) agrees that applying concepts to hypothetical cases in order to identify their extensions at various scenarios is part of a central method of philosophical theorizing. He is less inclined than we are, however, to think of knowledge that the results from deploying that method as a priori, on grounds that experience plays an ineliminable role in learning how to apply concepts to scenarios. As per our discussion of role 4, however, we think that even if experience does play such a role, this role is irrelevant to the a priority of knowledge acquired by using those methods. It is moreover worth noting that Williamson and we agree about the big picture, even if we disagree about details. He claims that "we must focus on the ways in which that knowledge [obtained by consideration of hypothetical scenarios] differs from both the stereotype of a *a priori* knowledge and from the stereotype of a *posteriori* knowledge" (190). Our discussion of 1–4 does what Williamson prescribes; namely, it addresses the ways in which knowledge acquired through consideration of hypothetical scenarios is a priori and ways in which such knowledge is a posteriori.

<sup>21</sup> Eklund (this volume) expresses sympathy with the claim that abduction is relevantly a priori, criticizing Hirsh (2009) and Hawthorne (2009) for presupposing that speculative, theoretical reasoning, including (for example) appeals to simplicity, delivers only a posteriori justification. Here we register that while Hirsh clearly presupposes that such reasoning delivers only a posteriori justification, Hawthorne may not. Hawthorne (2009) aims to show, against Hirsch, that metaphysicians' methods of belief formation are often on all fours with scientists' methods. This equity claim is consistent with metaphysicians' methods delivering a priori justification, provided that the relevant methods of scientists can do so. And elsewhere, Hawthorne expresses sympathy for the view that abduction can deliver a priori justification for belief in conditionals the antecedent of which describes an "experiential life history", and the consequent of which is whichever theory best explains some aspect of that life history (2002, 252); he also maintains that these conditionals are central to knowing metaphysical claims. For further discussion of the a priority of abduction, including how Hawthorne's view bears on this issue, see Biggs and Wilson (in progress b).

## 4.6 The Undermining of Carnap's Metaphysical Anti-realism

We have so far argued that a properly neo-Carnapian treatment of a posteriori necessities can implement the E2D strategy, understood as relying on an abductive epistemology of intensions. Call the E2D strategy, so understood, 'abductive two-dimensionalism'. Does a shift to abductive two-dimensionalism, as providing a new route to the a priori identification of necessary truths, have ramifications for other aspects of Carnap's philosophy? Yes. Most strikingly, it undermines Carnap's metaphysical anti-realism.

In 'Empiricism, Semantics, and Ontology', Carnap distinguishes two kinds of ontological question (or claim), expressible by appeal to the notion of a linguistic framework—a language with semantic rules sufficient for engaging in verificationistically acceptable discourse. On Carnap's account, the numbers framework, for example, is partly constituted by rules for proof-theoretically (analytically) confirming the existence of numbers (5, primes over 100, etc.), and the physical object framework is partly constituted by rules for empirically (synthetically) confirming the existence of physical objects (tables, electrons, etc.). Carnap's distinction between kinds of ontological questions is then cashed as a distinction between questions asked either 'internal' to some framework, or 'external' to any framework: internal questions have associated analytic or synthetic verification conditions, and so typically make sense; external questions do not have associated verification conditions, and so never make sense. Unlike mathematical or scientific questions, metaphysical questions are, Carnap claimed, paradigmatically external questions; hence his metaphysical anti-realism.

Though Carnap put his point in linguistic terms, the deeper source of his concern was his conviction that there are no appropriate standards of confirmation for metaphysical claims (see Wilson 2010). It is this supposed failure, after all, that prevents metaphysical questions/claims from being asked/asserted within a distinctively metaphysical framework. As such, assessing Carnap's metaphysical anti-realism requires attention not so much to semantic questions—pertaining, e.g., to whether there is a distinctively metaphysical quantifier, à la Sider (2009), Hirsch (2011), and others—but to whether metaphysical investigations have standards of confirmation sufficient to generally determine the outcome of metaphysical debate. We should start, then, by attending to Carnap's reasons for thinking not:

Suppose that one philosopher says: 'I believe that there are numbers as real entities.' [...] His nominalistic opponent replies: 'You are wrong: there are no numbers.' [...] I cannot think of any possible evidence that would be regarded as relevant by both philosophers, and therefore, if actually found, would decide the controversy or at least make one of the opposite theses more probable than the other. (1950, 56, 254)<sup>22</sup>

<sup>22</sup> On this traditional reading of Carnap he argues for metaphysical anti-realism primarily on epistemic grounds (specifically, verificationist or at least broadly empiricist grounds). Several contributors to this volume offer competing interpretations. While we discuss some of these in Appendix B, one should note that our interpretation remains standard, and thus, is dialectically apropos.

So Carnap reported. But was he correct? Suppose that one adopts abductive two-dimensionalism. One thereby accepts that abduction sometimes can confirm theories, and associated claims. So, if abduction can support some metaphysical claims over others, then plausibly abduction can confirm metaphysical claims. Going by what metaphysicians report, abduction *can* support some metaphysical claims over others. Hence Sider says, in characterizing ‘main-stream metaphysics’:

Competing positions are treated as tentative hypotheses about the world, and are assessed by a loose battery of criteria for theory choice. Match with ordinary usage and belief sometimes plays a role in this assessment, but typically not a dominant one. Theoretical insight, considerations of simplicity, integration with other domains (for instance science, logic, and philosophy of language), and so on, play important roles. (2009, 385)

Though not couched as such, this is a description of metaphysical deliberation as proceeding by way of abduction, i.e., by way of theory choice guided by attention to how well a given theory conforms to a range of broadly theoretical desiderata, which include “match with ordinary usage”, “considerations of simplicity”, “integration with other domains”, and other theoretical virtues. If this description of actual practice is broadly accurate, and we think it is (at least roughly), many metaphysicians rely on abduction and associated theoretical virtues, as supporting metaphysical claims. In that case, anyone who adopts abductive two-dimensionalism should accept that abduction can confirm metaphysical claims.

Now, one might be concerned about whether abduction is a properly empiricist mode of inference. Certainly there is a tradition, which at least superficially includes Hume, empiricism’s greatest defender, and which includes other self-identified empiricists unto the present day (notably, van Frassen, as per his 1980), according to which abduction is not a mode of inference in good standing. But most empiricists these days are typically happy to accept abduction—in part, because it seems impossible to do science without it. In any case, a neo-Carnapian who aims to reconcile Carnap’s intensional semantics with Kripke’s insights cannot take these concerns on board: as we’ve seen, only abduction has the resources to overcome widespread indeterminacy, and so provide a basis for a priori knowledge of a wide range of (conditional) modal truths.

Given that metaphysical claims can be confirmed, albeit defeasibly, by abduction, what prevents there from being a distinctively metaphysical linguistic framework? Nothing, by Carnap’s own lights. Carnap, or at least those neo-Carnapians aiming to accommodate a posteriori necessities by appeal to the E2D strategy, should allow that there is or in any case could be such a framework, from within which metaphysical questions can (could) be meaningfully and (like scientific questions, which are neither trivially true nor trivially false) substantively asked. But then, of course, Carnap’s case for metaphysical anti-realism, hinging as it does on the in-principle absence of a metaphysical framework and associated standards of confirmation, falls apart. The upshot is that the most natural post-Kripke version of Carnap’s intensional semantics undermines his metaphysical anti-realism.

One can react to this tension in pre- and post-Kripke Carnapian doctrine in a few different ways. One can backtrack, returning to a purely pragmatic interpretation of Carnap's methodology for the identification of intensions, perhaps on grounds that abduction cannot be an epistemic affair. As we discuss in Appendix A below, however, Carnap's motivations for a pragmatic interpretation of explication are unconvincing, and Kripke's results provide independent motivation for an epistemic interpretation of Carnap's methodology. Alternatively, one can insist that abduction has epistemic force when claims about intensions are at issue, but does not have such force when metaphysical claims are at issue. Since these different kinds of claim appear to be equally amenable to abduction, however, such a move would be ad hoc. Finally, one can embrace our result. This, we think, is the best option. After all, Carnap's metaphysical anti-realism was ultimately motivated by the worry that there is no substantive means of confirmation of metaphysical claims. Such a worry arises only if one ignores the possibility that abduction can be a warranted, if fallible, means of arriving at metaphysical results—just as it is a warranted, if fallible, means of arriving at results about intensions, scientific goings-on, other minds, and many other claims whose truth is not revealed by perception, conceivability, or any other form of comparatively direct access. Accordingly, we maintain that a post-Kripke neo-Carnapian should accommodate a posteriori necessities by endorsing abductive two-dimensionalism, and metaphysical anti-realism be damned.

## Appendix A: Epistemic vs. Pragmatic Interpretations of the Methodology of Intensions

Carnap maintains that we associate intensions with expressions on pragmatic rather than epistemic grounds. Why so? Carnap explicitly argues for his pragmatic understanding of explication only once:

In a problem of explication the datum, viz., the explicandum, is not given in exact terms; if it were, no explication would be necessary. Since the datum is inexact, the problem itself is not stated in exact terms; and yet we are asked to give an exact solution. This is one of the puzzling peculiarities of explication. It follows that, if a solution for a problem of explication is proposed, we cannot decide in an exact way whether it is right or wrong. Strictly speaking, the question whether the solution is right or wrong makes no good sense because there is no clear-cut answer. The question should rather be whether the proposed solution is satisfactory, whether it is more satisfactory than another one, and the like. (1950, 3–4)

This argument hinges on what it is to “decide in an exact way” whether a given explicatum *T* is “right or wrong” for a given explicandum *D*. If deciding in an exact way only requires having clear procedures for ranking competing explicata, then Carnap's method for associating concepts with intensions suggests that we *can* decide in an exact way whether *T* is right or wrong for *D*. If deciding in an exact way requires deciding with conclusive justification, then the fact that we cannot decide in an exact way whether *T* is right or wrong for *D* does not preclude (at least by Carnap's lights) our having justification for believing that *T* is right

or wrong for *D*; as Carnap says, “If by verification is meant a definitive and final establishment of truth, then no (synthetic) sentence is ever verifiable. We can only confirm a sentence more and more [...]” (1936, 420). So, Carnap’s only explicit argument for his pragmatic interpretation fails.

Is there an implicit motivation for a pragmatic interpretation anywhere in Carnap’s semantics? Kripke’s insights aside, yes. Carnap claims that explication consists in *replacing*, as opposed to *refining*, concepts—where replacement occurs when one exchanges one concept for another, and refinement occurs when one adjusts a concept without altering its identity, perhaps through mere precisification, perhaps by changing inessential elements.<sup>23</sup> Carnap then infers that there can be no fact of the matter about whether a candidate explicatum is correct for its explicandum, and thus, the choice is always a pragmatic one. Put another way, the argument is as follows: (1) explication is always replacement of one concept by a new concept; therefore, (2) there is never a fact of the matter about whether an explicatum is correct for its explicandum; therefore, (3) explication is not an epistemic enterprise.

Carnap’s rationale for (1) is revealed in an example. He claims that the “prescientific term ‘fish’ was meant in something like the sense of ‘animal living in water’; therefore its application to whales, etc., was entirely correct” (1950, 6). He also claims that using the prescientific concept (call it ‘Fish’) can be appropriate for certain purposes even for one who has acquired the zoological concept (call it ‘Fish\*’). Together, these claims suggest that any competent user of Fish who knows the relevant facts about sea animals (e.g., the evolutionary and deep biological facts) should still apply Fish to whales. If Fish\* merely refined Fish, then competent users who know the relevant facts should not apply Fish to whales. So, thinks Carnap, Fish\* replaces Fish—though ‘fish’ still expresses both concepts.<sup>24</sup> This discussion presupposes that *animal living in water* serves as a reference fixing description for our pre-scientific, natural kind concept of fish, as per traditional descriptivist theories of meaning. In turn, this presupposition implies that our pre-scientific concept includes whales in the extension of Fish, even for one who knows relevant biological/evolutionary facts.

Kripke’s rejection of this presupposition is, of course, at the heart of many of his insights about meaning and modality. Kripke compellingly argues both that descriptions that are a priori associable (in some loose sense) with natural kind concepts are typically not reference-fixing (at least not in the crude manner that traditional descriptivist theories presumed), and that natural kind concepts are typically not associated with an a priori reference-fixing description. Rather, he argues, natural kind concepts have a consistent extension all along, such that the ancient Greeks and we both have a concept of fish that excludes whales from its extension, even though only we are sufficiently informed about relevant facts to recognize that

<sup>23</sup> There are, of course, puzzles about how objects can persist through change, and those puzzles transfer readily to concepts. Rather than engage these puzzles here, we consider Carnap’s reasons for thinking that concepts cannot persist through changes that result from explication, showing that they are not compelling, even if they seemed compelling prior to Kripke’s work.

<sup>24</sup> This approach leaves us with Fish and Fish\* as distinct concepts in our conceptual repertoire. Carnap thinks that the use of Fish will diminish since Fish\* is more fruitful, which, after all, is why explication of Fish resulted in Fish\*. What holds for this explicandum (Fish) and its explicatum (Fish\*) holds more generally for any explicandum–explicatum pair, since fruitfulness always plays a role in identifying the explicatum, and fruitfulness always can lead to choosing an explicatum that is so dissimilar to its explicandum that there is relatively little classificational overlap—at least, so Carnap supposes.

exclusion.<sup>25</sup> So, the naive descriptivist theory required to motivate Carnap's assumption that explication involves replacement of one concept with another is simply outdated. As such, (1) is unmotivated; hence so is Carnap's rationale for thinking that explication is a pragmatic rather than an epistemic enterprise.

Had Carnap realized that his rationale for a pragmatic understanding of explication would be undermined, he might have been more open to an epistemic interpretation of abductive deliberation than one might expect. Strawson (1963) objects that explication, as Carnap conceives of it, is useless for philosophy:

[H]owever much or little [explication] is the right means of getting an idea into shape for us in the formal or empirical sciences, it seems *prima facie* evident that to offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of non-scientific discourse, is to do something utterly irrelevant—is a sheer misunderstanding, like offering a text-book on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart [...] laying down the rules of use of exact fruitful concepts in science [...] is not to solve the typical philosophical problems, but to change the subject. (504–6)

Strawson's objection presupposes that explication consists in replacing ordinary concepts. Carnap (1963) counters that he has "the impression that Strawson's view is based on the conception of a sharp separation, perhaps even a gap, between everyday concepts and scientific concepts. I see here no sharp boundary line but a continuous transition" (1963, 934). It is tempting to think of the "continuous transition" at issue as occurring *within* concepts. Carnap (1963) reinforces this temptation by claiming that the scientific concept of warmth is a "disambiguation" of the ordinary concept, which suggests that it is not, after all, a wholly new concept, but is rather a refinement of the prescientific explicandum—contrary to his earlier (1950, 8–15) discussion of the transition from the prescientific concept to the scientific concept of warmth, in which he suggested that explication involved replacement, again for broadly descriptivist reasons. Carnap's response to Strawson suggests that he might have been uncomfortable with thinking of explication in terms of replacement, even though his naïve descriptivism, as it manifests in (1)–(3), forced this result upon him. Perhaps, then, in light of Kripke's insights, Carnap would be open to thinking of explication as refinement.

Finally, one should note that thinking of explication as refinement, as Kripke's insights suggest we should, not only undercuts Carnap's best reason for interpreting semantics as a pragmatic enterprise, but also potentially directly motivates an epistemic interpretation. The idea, very roughly, is this: if a given explication involves refinement, then there *is* a fact of the matter about whether the explicatum at issue is correct for its explicandum, and thus, explication can be an epistemic enterprise. Unlike a pragmatic account of the assignment of intensions to expressions, then, an epistemic account is well-motivated. We find, then, that the most plausible post-Kripke updating of Carnap's intensional semantics combines a post-Kripke epistemic reading of Carnap's explication-based method for identifying intensions with the E2D strategy, resulting in abductive two-dimensionalism.

<sup>25</sup> *How* our concepts manage this is quite controversial. *That* our concepts manage this is not so controversial, thanks to Kripke. Accordingly, our reply to Carnap's argument for a pragmatic interpretation of semantics is dialectically apropos.

## Appendix B: Metaphysical Anti-realism as a Consequence of Verificationism

We assume a commonplace reading of Carnap's metaphysical anti-realism, according to which he advances this view on verificationist (or at least, broadly empiricist) grounds. There are competing interpretations of these motivations. Some of these are innocuous for our dialectical purposes—for example, a deflationist reading according to which metaphysical questions should be reinterpreted as internal questions having trivial answers (see also Levine, this volume, for a case against a deflationist reading). Other competing interpretations are less innocuous—in particular, interpretations that treat Carnap as neither an anti-realist nor a deflationist about metaphysics, and interpretations that divorce his attitude toward metaphysics from his verificationism. We consider certain such interpretations as forwarded by Thomasson, Kraut, and Sidelle (this volume); our points generalize to others who worry about the commonplace reading.

Thomasson (this volume) agrees with us that Carnap thinks of external metaphysical questions as misguided, but insists that understanding metaphysical questions as internal questions does not impugn ontological realism—after all, she thinks, to say that numbers exist according to a framework is to say that there really are numbers, in the only senses of “really” and “there are” that there are. She also claims that Carnap's rejection of external questions (qua epistemic questions) has little to do with his verificationism (empiricism) and much to do with his thinking that we cannot even use our terms (although we can mention them) when speaking externally.

Thomasson recognizes that this interpretation is strained. She claims merely that “*there is a way to interpret Carnap's view that does not rely on verificationism nor lead to anti-realism*”, not that Carnap advances such a view (our italics, 122). Accordingly, she sees her work as more “appropriation” than “historical interpretation” (124).

In any case, we can consider, as Thomasson does, whether Carnap *should have* grounded his attitude toward metaphysics in the belief that words cannot be used but only mentioned outside of a framework. Four considerations push towards a negative answer. First, it is not only metaphysical questions that cannot be asked by merely mentioning terms, it is any question at all: if an uttered word is merely mentioned, then nothing is said with it, regardless of whether the word would have metaphysical import if it were used. The appeal to the use-mention distinction, then, cannot explain why Carnap thinks of metaphysical questions as *especially* problematic. Second and relatedly, appeal to the use-mention distinction cannot explain why we cannot, according to Carnap, adopt a distinctively metaphysical framework (see Wilson 2010). Third, Carnap says that we can legitimately explore metaphysical questions (qua external questions) by treating them as pragmatic questions about which framework is best for one or another purpose, but metaphysical questions cannot be answered on pragmatic grounds if crucial terms are only mentioned. We understand how we might have pragmatic reasons to prefer Platonism to nominalism about this or that if Platonism and nominalism are at issue, but how could we have such reasons if ‘Platonism’ and ‘nominalism’, taken as uninterpreted words, are at issue? Fourth, insofar as we can have pragmatic reasons to prefer one framework to another—and Carnap clearly thinks we can—we can have epistemic reasons to prefer one framework to another if, as we maintain, the method that Carnap advances for making pragmatic decisions is epistemically significant.

Kraut (this volume) agrees with us that metaphysical questions are external questions, but insists that we should think of them as expressing our commitments to the pragmatic value of

our preferred frameworks, rather than as nonsensical ramblings. He admits, however, that his expressivist reading is nonstandard, and that the reading that we presume is “widely shared”; he offers several citations supporting or presuming our reading; and as such attributes the expressivist view at issue to ‘Carnap\*’, rather than to Carnap himself (31).

An expressivist reading of Carnap’s metaphysical anti-realism, moreover, does not threaten our dialectic. Kraut maintains that for Carnap\*, ontological claims are bound to explanatory considerations. Specifically, he thinks that Carnap\* treats existence claims as “expressions of...commitments to the explanatory ineliminability of a given discursive framework” (42). Consequently, those who disagree about an ontological claim should “make explicit (1) the data they seek to deal with; (2) their sense of what it would *be* to adequately deal with it; (3) their criteria for treating one way of dealing with it as superior to another” (40). So, for example, “arguments about the existence of a Judeo-Christian deity [understood as external, not internal] turn on disputes about best explanation of natural phenomena” (42). As such, on Kraut’s reading, like ours, the procedures for choosing among competing explanations rely on explication; and as such we can go on to ask: why do Carnap and Carnap\* think that the prescribed procedure tracks pragmatic but not epistemic value? For Carnap, we think, the answer is grounded in verificationist (empiricist) scruples. But Kraut treats Moore’s open question argument as decisive for Carnap\*, saying that since it *could be* pragmatically useful to adopt a way of talking that doesn’t track reality (if, say, number talk was useful even in the absence of numbers), the principles that establish the pragmatic value of a framework don’t establish the reality of the entities implied by that framework. But this is a decisive consideration only if reasons can’t be defeasible. We think, as most philosophers these days do, that reasons can be defeasible.

Finally, Sidelle (this volume) explores the connection between Carnap’s attitudes toward metaphysics and his verificationism more carefully than any other contributor. He concludes that Carnap’s clearest arguments against metaphysical theorizing rest on his verificationism (see especially 78–79). He also suggests, however, that Carnap offers some considerations which push towards anti-realism but do not presuppose verificationism. If Carnap’s metaphysical anti-realism can be grounded in considerations independent of any broadly empiricist epistemological scruples, our arguments would need to be adjusted. We cannot explore this issue adequately here, but here register that we doubt that these other considerations are independent of empiricist epistemological scruples.

We conclude that competing interpretations of the source and/or purport of Carnap’s metaphysical anti-realism either pose no pressing problems for our view, or are as yet unmotivated.

## References

- Beaney, Michael (2004). ‘Carnap’s Conception of Explication: From Frege to Husserl?’, in S. Awodey and C. Klein (eds.), *Carnap Brought Home: The View from Jena*. Chicago: Open Court Press.
- Biggs, Stephen (2011). ‘Abduction and Modality’. *Philosophy and Phenomenological Research* 83: 283–326.
- Biggs, Stephen and Jessica Wilson (forthcoming). ‘Conceptual Analysis and Abductive Deliberation’, *Synthese*.
- Biggs, Stephen and Jessica Wilson (in progress a). ‘Resolving Conceptual Indeterminacy from the Armchair’.

- Biggs, Stephen and Jessica Wilson (in progress b). 'The A Priority of Abduction.'
- Block, Ned and Robert Stalnaker (1999). 'Conceptual Analysis, Dualism, and the Explanatory Gap,' *Philosophical Review* 108: 1–46.
- Byrne, Alex, and James Pryor (2006). "Bad intensions," in M. Garcia-Carpintero and J. Macia (eds.), *Two-Dimensional Semantics: Foundations and Applications*. Oxford: Oxford University Press.
- Carnap, Rudolph (1936). 'Testability and Meaning,' *Philosophy of Science* 3: 419–71.
- Carnap, Rudolph (1947/1956). *Meaning and Necessity*, 2nd Edition. Chicago: University of Chicago Press.
- Carnap, Rudolph (1950). "Empiricism, Semantics, and Ontology," *Revue Internationale de Philosophie* 4: 20–40.
- Carnap, Rudolph (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.
- Carnap, Rudolph (1963). "P.F. Strawson on Linguistic Naturalism," in P.A. Schlipp (ed.), *The Philosophy of Rudolph Carnap*. La Salle, Illinois: Cambridge University Press.
- Chalmers, David and Frank Jackson (2001). "Conceptual Analysis and Reductive Explanation," *Philosophical Review* 110/3: 315–60.
- Chalmers, David (2006). "The Foundations of Two-Dimensional Semantics," in M. Garcia-Carpintero and J. Macià (eds.), *Two-Dimensional Semantics: Foundations and Applications*. New York: Oxford University Press.
- Chalmers, David (2006a). "Two Dimensional Semantics," in E. Lepore and B. Smith (eds.), *Oxford Handbook of the Philosophy of Language*. New York: Oxford University Press.
- Chomsky, Noam and Howard Lasnik (1993). "Principles and Parameters Theory," in *Syntax: An International Handbook of Contemporary Research*. Boston: Mouton De Gruyter.
- Feeney, Aidan and Evan Heit (2007). *Inductive Reasoning: Experimental, Computational, and Developmental Approaches*. Cambridge: Cambridge University Press.
- Gelman, S. A. and E. M. Markman (1986). "Categories and Induction in Young Children," *Cognition* 23: 183–209.
- Hawthorne, John (2009). "Superficialism in Ontology," in D. Chalmers and D. Manley (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. New York: Oxford University Press.
- Hirsh, Eli (2009). "Ontology and Alternative Languages," in D. Chalmers and D. Manley (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. New York: Oxford University Press.
- Hirsh, Eli (2011). *Quantifier Variance and Realism: Essays in Metaontology*. New York: Oxford University Press.
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- Kant, Immanuel (1781/1998). *Critique of Pure Reason*, in *Critique of Pure Reason*, trans. P. Guyer and A. Wood. New York: Cambridge University Press.
- Kripke, Saul (1972/1980). *Naming and Necessity*. Princeton: Princeton University Press.
- Lipton, Peter (1991/2004). *Inference to the Best Explanation*. New York: Routledge.
- Raffman, Diana (1994). "Vagueness without Paradox," *Philosophical Review* 103: 41–74.
- Schroeter, Laura (2004). "The Rationalist Foundations of Chalmers' 2D Semantics," *Philosophical Studies* 118: 227–55.

- Searle, John (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Sider, Theodore (2009). "Ontological Realism," in D. Chalmers and D. Manley (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. New York: Oxford University Press.
- Strawson, Peter (1963). "Carnap's View on Constructed Systems versus Natural Languages in Analytic Philosophy," in P.A. Schlipp (ed.), *The Philosophy of Rudolph Carnap*. La Salle, IL: Cambridge University Press.
- van Frassen, Bas (1980). *The Scientific Image*. New York: Oxford University Press.
- Williamson, Timothy (2007). *The Philosophy of Philosophy*, Malden, MA: Blackwell Publishing.
- Wilson, Jessica (2010). "Much Ado about 'Something': Critical Notice of *Metametaphysics*," *Analysis* 71: 172–88.
- Wilson, Mark (1982). "Predicate Meets Property," *The Philosophical Review* 91: 54f–89.
- Wilson, Mark (2006). *Wandering Significance: An Essay on Conceptual Behavior*. New York: Oxford University Press.