

# Two-Dimensionalism and the Social Character of Meaning

Derek Ball  
University of St Andrews

David Chalmers's two-dimensionalism purports to provide a framework for thinking about meaning and mental content that can provide substantial insight into rationality, modality, and the nature of mind. Such an ambitious program cannot but provoke opposition, and there have been a number of prominent critiques of two-dimensionalism. But much of this opposition is based on misunderstanding. A number of influential criticisms turn on interpreting two-dimensionalism as involving the view that the meanings of names can be given by descriptions associated with them by speakers.<sup>1</sup> But the best current version of two-dimensionalism – the view developed in Chalmers's recent work (Chalmers, 2002a, 2006, 2007) – is not descriptivist, and is not so easily refuted.

Two-dimensionalism is a serious contender, a view to be reckoned with. In Chalmers's hands, it promises to provide both an account of the cognitive significance of our thoughts, and a useful systematisation of the role of conceivability in modal epistemology, especially as it bears on Kripke/Putnam-style necessary a posteriori. Ultimately, however, I will argue that this promise is not fulfilled. Chalmers's development of 2D mimics Carnap's development of his own semantic views.<sup>2</sup> I maintain that Chalmers's Carnapian program faces serious difficulties with broadly Quinean objections. Quine argued that Carnap's program fails: it requires the existence of a priori truths, but (Quine claimed) no sentence is immune to revision, and hence that nothing is a priori. Quine's arguments for this claim can be resisted (Chalmers, 2011b). But – as I show – Burge-style scenarios in which a person changes her beliefs in response to empirical evidence (especially testimony from experts in her linguistic community) provide revision scenarios that are much harder to resist, particularly from the two-dimensionalist's point of view. I argue that it is possible that there

---

<sup>1</sup>I have in mind many of the arguments in Soames (2007); Schiffer (2003); Byrne and Pryor (2006). The critics have legitimate targets in Jackson (1998) and perhaps Chalmers (1996), but I will not consider descriptivist variants of 2D further. As will become clear, other critics of 2D have come close to anticipating my own objection; see my discussion of Block and Stalnaker (1999) and of Yablo (2002) in section 1.1. In my view, these criticisms can be answered (and indeed, have been answered in Chalmers (2006, 2007)). A further source of confusion is the idea that two-dimensionalism is an attempt to provide a compositional semantics for natural language. Of course, techniques derived from two-dimensional modal logics have been profitably applied to natural language by Kaplan (1977) and others, and I have no objection to these semantic programs. It is even possible that 2D will play a role in giving a semantics of belief reports (see Chalmers (2011a)). But the distinctive epistemic theses of Chalmers's two-dimensionalism have very little to do with natural language semantics, as it is understood by (say) Heim and Kratzer (1998); for one thing, Chalmers's primary intensions can vary between tokens of a single natural language word type (see section 2.2 below).

<sup>2</sup>These Carnapian roots are especially emphasised in Chalmers (2013).

be a Quinean thinker who is willing in principle to revise any of her beliefs in response to empirical evidence, especially testimony; that for such a thinker, nothing would be apriori by 2D lights; and that this possibility is enough to undermine the 2D position.<sup>3</sup>

I begin by developing two-dimensionalism, and especially the 2D account of cognitive significance, with special attention to Chalmers's response to plausible objections from Block and Stalnaker and from Yablo. In section 2, I consider two ways of developing the 2D account of cognitive significance, and argue that the possibility of a Quinean thinker who is very similar to us is enough to refute each of them. The third section considers the application of the 2D apparatus to modal epistemology; here, too, I argue that the Quinean considerations show that 2D gives little insight.

## 1 Two-Dimensionalism Motivated

Chalmers's 2D is motivated by a traditional picture of meaning on which apriority and necessity coincide. This picture can be developed in a broadly Carnapian framework (Carnap, 1956). On this view, the meanings of sentences are (or determine) *intensions*, functions from possible worlds to truth values (or, equivalently, sets of possible worlds). We begin with the set of all the worlds: this is the necessary proposition, the proposition that is true in every world. It is also the proposition that is knowable apriori, since we can know apriori that we are in *some* world, even if we cannot know which one. The goal of empirical enquiry is to rule out worlds as non-actual; empirical information tells us which world we are in.

Chalmers's view is designed to preserve the intuitive appeal of the Carnap picture, while permitting an adequate account of challenging examples first presented by Saul Kripke (1980). Kripke claimed that there are both contingent apriori and necessary aposteriori truths. For example, the claim that *if there is a unique tallest person alive, then the actual tallest person alive is the tallest person alive* seems to be knowable apriori, yet it is contingent, since the actual tallest person might not have been the tallest. By contrast, the claim that *Hesperus is identical to Phosphorus* is necessary if true, but seems to be knowable only aposteriori. Any theory that holds that the necessary is coextensive with the apriori must provide an alternative account of Kripke's examples.

Chalmers addresses the Kripkean problem by claiming that representations<sup>4</sup> have two types of content: *primary intensions*, which are associated with apriority and aposteriority,

---

<sup>3</sup>In section 2.4, I suggest further that the objections apply to any view that attempts to account for the cognitive significance of our thoughts in terms of a notion of the apriori that is empirically indefeasible. 2D is perhaps the most prominent account of this type in the current literature, but the objections presented would generalise to any such account.

<sup>4</sup>The two-dimensionalist story is in principle applicable both to sentences and to mental representations such as beliefs. I use 'representation' to cover both linguistic and mental representations.

and *secondary intensions*, which are associated with necessity and possibility in the sense relevant to Kripke's cases (i.e., 'metaphysical' necessity and possibility). Thus no single content is both necessary and aposteriori, nor is any single content both contingent and apriori. In particular, Chalmers defends the following view:

**The Core Thesis** A token representation  $S$  is a priori (epistemically necessary) iff the primary intension of  $S$  is true at all worlds (i.e., metaphysically necessary). (Chalmers (2006, p. 64), Chalmers (2007, §3.1))

### 1.1 Carnap and Quine (and Block and Stalnaker)

Quine (1953) famously argued that semantic projects like Carnap's depend on notions like analyticity and synonymy, but that there is no way to provide an acceptable account of these notions. In response, Carnap (1955) and Mates (1951) offered a procedure for determining an expression's intension: intensions can be determined by asking subjects not only about the actual extension of terms, but about what the term's extensions *would be* in counterfactual scenarios. Two expressions are synonymous if they have the same intension; a sentence is analytic if its intension is true at every world.

An analogous debate has taken place regarding primary intensions. Block and Stalnaker (1999) and Yablo (2002) played the role of Quine, alleging that Chalmers cannot provide a way of determining an expression's primary intension that can satisfy the Core Thesis. For example, Block and Stalnaker argue that we cannot simply consider the semantic values a given expression *would have* in various possible worlds; since any expression could have meant anything at all, the intensions thus generated would be uninteresting (Block and Stalnaker, 1999, §10). Chalmers agrees, noting that 'there are worlds in which the string 'bachelors are unmarried' means that horses are cows,' and thus that on such a view, 'the Core Thesis is obviously false' (Chalmers, 2006, p. 67). Similarly, Yablo (2002, pp. 470-1) claims that it is knowable apriori that if 'sibling' means number, then sisters are not siblings. With this in mind, Yablo invites us to consider a world in which 'sibling' means number. Given the allegedly apriori conditional, in such a world, sisters are not siblings. But this entails that it is not apriori that sisters are siblings. Since the same form of argument could be applied to any putatively apriori truth, Yablo concludes that the strategy 'drain[s] the class of conceptual necessities of all its members' (2002, p. 471).

Chalmers responds by providing a procedure that improves on Carnap's. On Chalmers's view, the value of a primary intension is not determined by properties of counterfactual tokens. Instead, primary intensions 'turn on the epistemic properties of an expression in the actual world' (Chalmers (2007, §5), see also Chalmers (2006, §3.1)). Chalmers claims that

‘once a subject is given enough information about the character of the actual world, then they are in a position to make rational judgments about what their expressions refer to and whether their utterances are true’ (Chalmers (2007, §3.4), see also Chalmers (2006, §3.6)). This (purported) ability allows us to determine a representation’s primary intension in roughly the following way: given the actual epistemic properties of a representation, and sufficient information about a possible world, we can determine what the extension of the representation would be if that world is actual. The primary intension is the function from worlds to extensions generated by considering each world as actual (that is, by considering what would be true in each world, on the assumption that that world is actual). A representation is apriori if its primary intension is true at every world.<sup>5</sup>

Of course, we humans lack the time and memory necessary to entertain comprehensive descriptions of an entire world. Nonetheless, it is plausible that there are meaningful sentences of our language that we could evaluate in worlds considered as actual only by entertaining comprehensive descriptions of those worlds. Similarly, we lack the intelligence necessary to evaluate many meaningful sentences of our language (for example, highly complex logical or mathematical claims). In order to assign appropriate primary intensions to these sentences, we must idealise, setting aside our contingent limitations and dispositions to err. So primary intension is determined by our rationally idealised judgments about the extensions of our representations (given their actual epistemic properties) in worlds considered as actual.

It is clear that the intension thus generated need not be captureable by any finitely statable descriptive sentence of a natural language. This fact blunts the force of standard criticisms of two-dimensionalism.<sup>6</sup> For example, Schiffer (2003, p. 148), complains against 2D that in many cases, no uniquely identifying description will be psychologically available to ordinary speakers. Similarly, Byrne and Pryor (2006, §§4-6) argue that Kripke’s Gödel-Schmidt case shows that properties associated with a name by speakers, such as *being the discoverer of the incompleteness of arithmetic*, do not fix the reference of that name. But of course Chalmers does not claim otherwise. What fixes the reference is speakers’ rational judgments about the application the name in various cases considered as actual. These judgments may or (more typically) may not be guided by properties that are easily describable and readily available to reflection. So Kripkean anti-descriptivist arguments of

---

<sup>5</sup>My discussion abstracts away from some of the details (e.g., the construction of scenarios and the notion of a canonical description of a world) presented in Chalmers (2006); these details do not affect the criticism of 2D that I develop below.

<sup>6</sup>Of course, not every criticism turns on a descriptivist interpretation; in particular, the objections developed in series of papers by Laura Schroeter (2003; 2004; 2005) are congenial to my own view.

the sort developed by Schiffer and by Byrne and Pryor have little force against Chalmers's theory.

## 1.2 The 2D Account of Cognitive Significance

Block and Stalnaker's objection plays a role in an argument designed to show that primary intensions are not even well-defined. Perhaps this style of objection could be pressed further, but it is plausible that our rationally idealised judgments about the application of our representations with their actual epistemic properties do determine an intension, and hence that Chalmers has successfully defined a relation to a content. But this in itself is hardly an achievement; we can easily define indefinitely many relations to contents. (For example, I hereby introduce the term *z-content* with the stipulation that all beliefs have the *z-content* that the author of this paper is a genius.) Of course, such 'contents' are to be ignored unless they can play a role in some psychological theory. Relations to propositions are cheap. They deserve our attention only if they do some work. The crucial question, then, is this: what work can primary intensions do?

Chalmers claims that primary intensions provide a 'constitutive connection between meaning, reason, and modality', and that they therefore promise to play a key role in explaining a thought's cognitive significance and thus providing an analogue of Fregean sense (Chalmers, 2006, p. 64). (Chalmers also maintains that primary intensions vindicate a deep link between conceivability and metaphysical possibility, despite Kripke/Putnam-style examples of the necessary a posteriori; I return to this motivation for 2D below.) The simplest sort of case is Frege's contrast between the cognitively significant "Hesperus = Phosphorus" and the cognitively trivial "Hesperus = Hesperus". On the 2D account, this contrast is explained in terms of necessity of primary intension: the claim is that "Hesperus = Hesperus" is true in every world considered as actual, and hence has a necessary primary intension, while "Hesperus = Phosphorus" is not and does not. The account can be extended to explain closely related contrasts in reasoning:

**Game Show** Hammurabi and Shammurabi are contestants on a game show. Each will be awarded a desired prize if he is able to name a bright object visible in the evening sky. Each believes that Hesperus is visible in the evening sky. Hammurabi also believes that Hesperus is bright. Shammurabi believes only that Phosphorus is bright; he does not believe that Hesperus is bright. Hammurabi names Hesperus and wins the prize; Shammurabi goes home empty handed.

What explains the contestants' behaviour? The crucial fact is that Hammurabi is able to perform the following inference:

**Inference A**

- (1) Hesperus is bright.
- (2) Hesperus is visible in the evening sky.
- (C) Hesperus is both bright and visible in the evening sky.

Shammurabi, on the other hand, is unable to infer (C) from (1) and (2), since he does not accept (1). He does accept (3), but nonetheless fails to perform Inference B:

**Inference B**

- (3) Phosphorus is bright.
- (2) Hesperus is visible in the evening sky.
- (C) Hesperus is both bright and visible in the evening sky.

This sort of case is, of course, familiar. The proponent of 2D attempts to account for it in terms of primary intension ([Chalmers, 2006, 2002b](#), p. 106). In inference A (as performed by Hammurabi), every world in which the primary intensions of the premises is true is a world in which the primary intension of the conclusion is true; in inference B (as entertained by Shammurabi), on the other hand, there are worlds – roughly, worlds in which “Hesperus” refers to something different than “Phosphorus” – in which the premises are true but the conclusion is false. Inference A is therefore rationally conclusive for Hammurabi, while inference B is not rationally conclusive for Shammurabi. Since Hammurabi and Shammurabi are rational, this can help explain why Hammurabi makes inference A, while Shammurabi does not make inference B.

In short, the 2D explanation of the distinction between “Hesperus = Hesperus” and “Hesperus = Phosphorus”, and of the contrast between Hammurabi and Shammurabi in the described scenario, turns on the following principle:

- Cognitive Significance Principle**
1. A representation is cognitively significant if and only if it has a contingent primary intension (i.e., is not apriori), and is cognitively trivial otherwise.
  2. An inference is rationally conclusive if and only if every world in which the primary intensions of the premises are true is a world in which the primary intension of the conclusion is true, and is rationally inconclusive otherwise.

I will begin my discussion in the next section by focusing on the account of the Fregean distinction between cognitively significant (e.g., “Hesperus = Phosphorus”) and cognitively trivial (e.g., “Hesperus = Hesperus”) thoughts; as the discussion progresses, I show how my arguments can be extended to cover the account of rational transitions in thought as well.

### 1.3 What Sort of Account is 2D?

If this account can be made to work, it would be attractive. But exactly what sort of account is it? Chalmers (and most of the literature on 2D) is inexplicit, and there are at least two importantly different possibilities:<sup>7</sup>

1. The 2D account might be part of a *philosophical analysis* or another sort of *metaphysically necessary* philosophical account of cognitive significance.<sup>8</sup>
2. The 2D account might be a metaphysically contingent generalisation that plays a role in a (perhaps causal) *explanation* of cognitive significance, or in psychological explanations of behaviour that depend on this phenomenon.

It is important to be clear about what sort of account 2D offers, since this determines what sort of arguments against it will be appropriate. In particular:

1. If the 2D account is intended as an analysis or some other sort of metaphysically necessary account, then the mere possibility of a thinker who exhibits a difference in cognitive significance between “Hesperus = Hesperus” and “Hesperus = Phosphorus” (of the sort that Frege described), but who does not satisfy the 2D account, is a counterexample sufficient to refute the account.
2. If the 2D account is a part of an *explanation*, it would not be refuted by the mere possibility of such a thinker. Nonetheless, it is often possible to refute a claimed explanation by attending to merely possible cases. This is because explanations must

---

<sup>7</sup>A third possibility is that the 2D account provides a *sufficient condition* for the cognitive significance phenomena to occur, which would play a role in explaining *how it is possible* for such phenomena to occur. (Such “how possible” explanations have important precedents in philosophy; see the introduction to Nozick (1981) for discussion.) A recent model of this sort of theory is Jerry Fodor’s theory of content, which attempts only to provide ‘naturally specifiable [...] sufficient conditions for a physical state to have an intentional content.’ (1990, p. 96.) But I set this aside for two reasons. First, it is pretty clear that this is not what Chalmers has in mind; he surely intends his account to explain our actual thinking and reasoning. And second, there is no real mystery about how it is possible for the cognitive significance phenomena to occur, once the resources needed to generate primary intensions are in place. Primary intensions depend on the ability to entertain detailed descriptions of possible cases and to make judgments about them; it is not surprising that the cognitive resources necessary for this ability can provide a sufficient condition for the cognitive significance phenomena. (This contrasts with Fodor’s theory of content; many philosophers have found it mysterious how intentionality could arise from the purely naturalistic resources to which Fodor appeals.)

<sup>8</sup>This interpretation is suggested by some of Chalmers’s writing: e.g., Chalmers (2006, p. 105, 108)

isolate the *difference-makers* that the explanandum *depends on*. Roughly, an explanation must be such that without the explanans, the explanandum does not occur. Different theories of explanation have different ways of implementing this idea. On Scriven's view (2008), the explanans must be a part of a (certain sort of) valid deductive argument that entails the explanandum, and (crucially) it must be an *essential* part of the argument: if the explanans is removed from the argument, the remaining premises must fail to entail the conclusion. On Woodward and Hitchcock's view (2003), an explanation must reveal certain patterns of counterfactual dependence, thus enabling one to answer a variety of *what-if-things-had-been-different* questions. But the most straightforward test is the simpler counterfactual:

**Counterfactual Test** E1 explains E2 only if: if E1 were not the case, E2 would not be the case.

Of course, this is not a complete account of explanation; but even proponents of other views typically admit that it is an accurate diagnostic in a range of cases (e.g., Scriven 2008, p. 111). Since the cases where it is most problematic (such as cases of preemption (see Lewis (1973) and Schaffer (2000) for discussion)) will be irrelevant to our discussion, I will appeal to it as a rough-and-ready test. But I will also appeal to other marks of a good explanation. In particular, a good explanation provides materials that can be used to explain similar cases. If phenomena very similar to the explanandum are observed in cases that the explanans cannot be adapted to cover, this suggests that the genuinely explanatory factors are being missed. A better explanation could be adapted to cover these phenomena as well.<sup>9</sup> (This is a core insight of unificationist views of explanation, such as Kitcher (1989).)

To see how these tests can be applied, consider an example of an inadequate attempt at explanation:<sup>10</sup>

- (a) The object that I dropped is an igneous rock.
- (b) Igneous rocks dropped near the Earth fall.
- (c) The object that I dropped fell.

---

<sup>9</sup>This style of argument is familiar in the literature on externalism and narrow content. Proponents of narrow content have sometimes argued that psychological explanations should cover (for example) both my twin-Earth doppelgänger and I. Externalists have also availed themselves of this sort of argument; for example, Williamson (2000, p. 62) argues that knowledge is explanatorily preferable to true belief because it covers other closely related cases.

<sup>10</sup>Are such attempts are not explanations at all, or are they merely bad explanations? Resolving this terminological issue is of little importance.



Intuitively, the problem with this attempted explanation is that it focuses on specific facts about the case – in particular, that the object in question is an igneous rock – that are irrelevant. That the rock is igneous does not make a difference; it is not what the falling depends on. The counterfactual test is one mark of this: it is not the case that if the object I dropped had not been an igneous rock, it would not have fallen. Alternatively, we can see that there are a variety of (actual and possible) similar cases in which a dropped object that is not an igneous rock falls. A better explanation – appealing, say, to the fact that the object I dropped has mass and the law of gravitation – can be adapted to cover these cases as well.

Similarly, even if the 2D account is intended as a part of a contingent explanation, it still must cover both actual and nearby possible cases. Thus if we exhibit the phenomena described in section 1.2, but do not satisfy the 2D account, this refutes the account construed as an explanation; and if possible thinkers very much like us exhibit the phenomena described in section 1.2 in just the way we do, but do not satisfy the 2D account, this is strong evidence that the account is a flawed explanation.

In the next section, I show how each of these interpretations of 2D can be criticised. I begin by returning to Quine’s critique of Carnap.

## 2 Two-Dimensionalism Criticised

### 2.1 The Quinean Argument

In what follows, I focus on a particular consequence of the 2D account of cognitive significance. The account rests on the Cognitive Significance Principle, the first part of which states that a thought is cognitively significant if and only if it has a necessary primary intension. This entails:

**2D Rationalism (2DR)** Every thinker who distinguishes between cognitively significant and cognitively trivial thoughts has some thoughts with necessary primary intensions.

Given the Core Thesis, (2DR) amounts to the claim that for a certain sort of thinker, some thoughts are apriori. This should bring to mind another of Quine’s attacks on Carnap. Quine maintained that no token sentence or belief is apriori, because empirical evidence can tell against any of them:

[N]o statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between such a shift and the shift

whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle?  
 (Quine (1953, p.43); cf. also Neurath (1959, esp. p. 204))

Quine's talk of revision suggests change in view over *time*. But we can reconstruct the point in terms of difference of view across *counterfactual scenario*, so as to bear more directly on Carnap's proposal for determining intension. Suppose that we are trying to determine the intension of some statement *s*. Recall that the suggested procedure for determining intensions is to solicit judgments about the extensions of their terms would be in various scenarios. If Quine is right, then there is some (possibly counterfactual) scenario in which we reject *s*. So we need only ask our subjects about this scenario; they will reply that *s* is false in it. Thus *s* will not be necessary (since the intension of *s* is false at some world). Thus if Quine is right, Carnap cannot maintain that any statements are necessarily true. The same style of argument can be adapted to Chalmers's more sophisticated story about primary intensions; so if Quine is right, then none of our representations has a necessary primary intension, (2DR) is false, and Chalmers's explanation of cognitive significance fails.

Of course, many have found Quine's arguments for the claim that none of our actual beliefs are apriori unconvincing. But recall the discussion of the previous section. If 2D is construed as a part of an *analysis* or other metaphysically necessary account, then it must apply to all possible thinkers who exhibit the relevant phenomena. So on this interpretation, (2DR) must be a necessary truth. Let a *Quinean thinker* be a rational thinker for whom no token belief is immune to revision; that is, in 2D terms, a thinker for whom no token belief has a necessary primary intension. If such a thinker is *even possible*, then – assuming that she would continue to distinguish between cognitively significant and cognitively trivial thoughts in just the way we do – (2DR) is not a necessary truth, and the 2D account cannot be an analysis.

The mere possibility of a Quinean thinker is not enough to refute the 2D account construed as an explanation of the cognitive significance phenomena. But recall that a good explanation must pick out the factors that are difference-makers – the factors on which the explanandum depends – and that we can test for such dependence by considering the counterfactual: if the explanans had not obtained, the explanandum would not have obtained. Relatedly, if there are very similar cases in which the explanans is absent but the explanandum still occurs, this suggests that the explanatory generalisation is being missed. In the case of the distinction between cognitively significant and cognitively trivial thoughts, the explanans is that the cognitively trivial thoughts are apriori. If it is construed as an explanation, then, the 2D account is committed to the following:

**2D Explanation Commitment (2EC)** 1. If we were Quinean thinkers, we would not

distinguish cognitively significant and cognitively trivial thoughts in just the way we now do; and

2. it is not possible that there be Quinean thinkers who differ from us only in small and subtle ways and who continue to distinguish cognitively significant and cognitively trivial thoughts in just the way we now do.

This, then, is the Quinean master argument against the 2D account of cognitive significance:<sup>11</sup>

1. If primary intensions are an interesting form of narrow content, then the 2D account of cognitive significance is either:
  - (a) an *analysis* or other *metaphysically necessary philosophical account*; or
  - (b) a (possibly contingent, e.g. causal) *explanation*.
2. The 2D account entails (2DR).
3. If (1a), then (2DR) must be a metaphysically necessary truth. (From (2).)
4. But Quinean thinkers who distinguish cognitively significant and cognitively trivial thoughts are possible, so (2DR) is not metaphysically necessary.
5. So, the 2D is not an adequate analysis or metaphysically necessary account (from (3), (4)).
6. If (1b), then (2EC) is true.
7. But either we are Quinean thinkers, or small and subtle changes to our psychology would turn us in to Quinean thinkers (who would continue to distinguish cognitively significant and cognitively trivial thoughts in just the same way). So (2EC) is false.
8. So, 2D is not an adequate explanation (from (6), (7)).
9. So, primary intensions are not an interesting form of narrow content (from (1), (5), (8)).

The crucial undefended premises are (4) and (7). In the next two sections, then, I defend the following claims:

---

<sup>11</sup>Note that the argument does not purport to show that primary intensions would not be defined for a Quinean thinker; I am granting this, at least for the sake of argument. Rather, the objection is that if a Quinean thinker is possible, then primary intensions cannot be the basis of a successful account of cognitive significance.

1. A Quinean thinker is *possible*. This shows that (2DR) is not a necessary truth, and hence refutes the 2D account of cognitive significance construed as an analysis.
2. We are Quinean thinkers or are very much like Quinean thinkers. Since the 2D account of cognitive significance would not apply to a Quinean thinker, it is not a good explanation of our own reasoning.

## 2.2 The Possibility of A Quinean Thinker: Burge's Cases

Primary intensions are determined by our dispositions to make judgments about possible cases under rational idealisation. It is reasonably obvious that there could be a thinker with broadly Quinean dispositions; for example, there could be a thinker who is disposed to reject logical judgments (including, perhaps, judgments of the form  $A=A$ ) in response to evidence about quantum mechanics. But it is not obvious that these dispositions could remain *under rational idealisation*. (One version of this point is developed by Chalmers (2011b).) But we can sidestep this kind of worry. I want to argue that *the procedure that Chalmers describes for generating primary intensions – and in particular, his account of Burge-style cases – ensures that there could be Quinean thinkers*.

Recall the famous cases that Burge (1979; 1986) used to support externalism about the mind. Consider Bert, Burge's character who has a belief that he would express with the sentence, 'Arthritis can occur in the thigh'. Bert is willing to accept correction on this matter from those he takes to be more expert than himself. In particular, when a doctor tells him that arthritis is a disease of the joints, he revises his belief, perhaps replying: 'I always thought that arthritis could occur in the thigh. What a silly mistake!'

Burge argues that in this sort of case, Bert uses the word 'arthritis' with the same meaning as his doctor, and possesses the same concept of arthritis as his doctor. Two sorts of consideration are particularly telling. Suppose that Bert and the doctor do not possess the same concept of arthritis. Then they share no arthritis-related beliefs, and neither agree nor disagree about matters arthritic. But surely Bert shares with his doctor the belief that arthritis is a disease, while also doubting another of his doctor's beliefs, that arthritis occurs only in joints. (Bert can even share these beliefs with speakers of other languages, where the agreement could not be merely metalinguistic.) Moreover, if Bert and his doctor are not using words with the same meaning or thinking with the same concepts, then it is hard to see why Bert accepts the doctor's correction (rather than responding that the doctor is quibbling about semantics rather than addressing his concern about the disease in his thigh), and later reports, 'I used to believe that arthritis could occur in the thigh. What a silly mistake!' (since if social phenomena are irrelevant to mental content, he had no such

belief and made no such mistake).

Burge's cases are closely related to the Kripke/Putnam cases of the necessary aposteriori that 2D is designed to explain away: in both cases, there is a necessary truth that a subject can know only by empirical means – scientific investigation in the case of Kripke's and Putnam's cases, testimony in the case of Burge's cases.<sup>12</sup> Chalmers therefore claims that 2D offers an illuminating account of the Burge cases:

Here, the crucial factor is that Bert uses the term 'arthritis' with semantic deference, intending (at least tacitly) to use the word for the same phenomenon for which others in the community use it. We might say that this term expresses a deferential concept for Bert: one whose extension depends on the way the corresponding term is used in a subject's linguistic community. It is clear that for deferential concepts, extension can depend on a subject's environment, as can subjunctive [i.e., secondary] intension. (Chalmers, 2002a, §6)

This account of the Burge phenomena exploits two features of Chalmers's view. First, there is the idea that primary intensions are properties of representation tokens, so that different speakers might differ with respect to the primary intension that they associate with a given public language expression. This allows Chalmers to maintain that a claim like 'Arthritis is a disease of the joints' could be apriori for some (non-deferential) speakers, but aposteriori for others who (like Bert) are willing to accept expert correction. Second, Chalmers maintains that the truth of belief reports like 'Bert and his doctor both believe that arthritis is a disease' and presumably also agreement reports like 'Bert and his doctor agree that arthritis is a disease' does not require that Bert and his doctor believe that arthritis is a disease under a common primary intension. Instead, Chalmers claims that the truth of such reports requires only that Bert and his doctor have a belief with the same *secondary* intension (i.e., the secondary intension of 'arthritis is a disease' as used in the report), and that the primary intensions of their beliefs bear a certain 'coordination relation' to the primary intension of 'arthritis is a disease' as used in the report (Chalmers, 2011a). This framework is difficult to evaluate without a detailed account of the coordination relation, but at least it has the structure necessary to accommodate the data that drive the Burge arguments while also maintaining a sense in which content does not depend on the social environment.

The Burgean considerations are in some respects similar to the sort of objection discussed by Block and Stalnaker and by Yablo, in that both involve scenarios in which linguistic usage

---

<sup>12</sup>On some views, knowledge gained by testimony may count as apriori (e.g., Burge (1993)). But this sort of view stands in serious tension with the Core Thesis and the account of primary intension described above.

is different. Chalmers's response to this sort of objection is that the two-dimensionalist can dismiss these cases because they involve change of meaning (in the sense that in these relevant scenarios, terms are not used with their actual epistemic properties). There is certainly a sense in which change of meaning is also relevant to Burge's cases. For example, Burge's two worlds differ in that in one, 'arthritis' is used exclusively to apply to diseases of the joints, and in the other, 'arthritis' is used also to apply to disease of the thigh. There is a difference in the extension of 'arthritis' between the two worlds, and on the 2D view, a corresponding difference in secondary intension; according to one usage, 'Arthritis occurs exclusively in joints' expresses a metaphysically necessary truth, while according to the other usage it expresses a contingent falsehood. But it would be very hard for the two-dimensionalist to avail herself of the claim that there is a difference in *primary* intension. For the strategy for resisting Block and Stalnaker was to consider the representations as they are actually used, with their actual epistemic properties. And it is clear that (at least for many normal speakers) many of our representations are actually epistemically responsive to information from experts in something like the way that Burge claimed. When we consider the extensions of these representations in worlds considered as actual, holding their epistemic properties fixed, we judge that their extensions correspond to the uses of expert speakers in that community.

It is a consequence of Chalmers's position that a person who possesses a concept 'deferentially' can know few truths involving that concept apriori, since it is epistemically possible for such a person that other members of the linguistic community will contradict her beliefs involving the concept. For example, no armchair reasoning can put Bert in a position conclusively to rule out the possibility that he is in a world in which experts take the sentence, 'Arthritis can occur in the thigh' to express a falsehood. It is therefore epistemically possible for Bert that his belief is false. If primary intensions satisfy the Core Thesis, this means that this belief is not apriori for Bert. And the more deferential a person is with a given concept, the fewer apriori truths involving that concept there will be for him. For example, probably many speakers could be convinced by expert physicians to give up the claim that arthritis is a disease, or that it occurs in the joints. (Perhaps it is a part of some natural healing process, or perhaps it is a mere symptom of some other disease; and perhaps doctors use 'arthritis' to pick out a phenomenon that causes joint pain while itself occurring outside of joints.) For these speakers, it would not be apriori that arthritis is a disease, or that it occurs in joints.

I argued above that there is a problem for 2D if it is possible that there be some being who distinguishes cognitively significant and cognitively trivial thoughts in just the way we

do, for whom no thought is apriori. Now it is an empirical question to what extent we are willing to accept the testimony of experts. Burge points out that Bert could have refused to accept the doctor's correction and insisted that as he is using the term, he does have 'arthritis'. In this case, Burge's arguments do not seem to apply. And it is possible that some speakers are unwilling to defer to experts with respect to very many beliefs. But surely it is also possible that there be some being who is *more* willing to accept correction than we are, and indeed is willing to accept correction with respect to *any* of her beliefs. To put the point more precisely: when we consider as actual scenarios in which the usage of our terms differs, we judge in at least some cases that the extensions of our terms in those worlds are determined in part by testimony from experts in the community. These judgments would survive rational idealisation (a point I return to below). So some of our concepts are what Chalmers calls 'deferential'. I want to argue that there could be a thinker who (even given rational idealisation) is disposed in principle to defer to community usage with respect to all of her representations.

It only takes a little thought to formulate similar cases for representations that have traditionally been held as paradigms of the apriori:

**Matters of 'definition'** Many of the classic Burge cases are examples of truths that are in some sense a matter of definition. And it certainly does seem that most speakers are willing to revise their beliefs on such matters when faced with suitable numbers of suitably expert speakers. A pertinent case: most philosophers are under the misimpression (perpetuated by Burge and by my discussion so far) that arthritis is a disease of the joints. But rheumatoid arthritis can spread outside of joints, and can occur in the thigh.<sup>13</sup> Thus, contrary to widespread philosophical rumor, arthritis can occur in the thigh after all. Other cases are easy to generate. For example, we could easily have used the term 'bachelor' to pick out all and only unmarried men under the age of 65. Surely it is conceivable that we discover that this is the common usage among other speakers, the meaning given in dictionaries, and so forth; and surely many ordinary speakers would defer in this matter.

---

<sup>13</sup>According to The U.S. Department of Health and Human Services/National Institutes of Health's publication *Handout on Health: Rheumatoid Arthritis*: 'Rheumatoid arthritis is an inflammatory disease that causes pain, swelling, stiffness, and loss of function in the joints. It has several special features that make it different from other kinds of arthritis. [...] It can also affect other parts of the body besides the joints' (2004, p. 1). (Remarkably, this can include 'inflammation of the blood vessels, the lining of the lungs, or the sac enclosing the heart' (2004, p. 5).) This fact can also be found in standard medical reference books, e.g. Macpherson (2002): 'Rheumatoid Arthritis: A chronic inflammation of the synovial lining of several joints, *tendon sheaths or bursae* which is not due to sepsis or a reaction to uric acid crystals' (p. 541, my emphasis).

**Mathematical truths** Mathematical beliefs, too, can be overturned on Burge-style grounds.

For example, most of us would be willing to defer to a mathematician as to whether 1 is prime. For example: before writing this paper, I had a well-justified belief that 1 is not prime. Nonetheless, in the course of writing, I decided to double-check in a math book. I was prepared to defer.<sup>14</sup> This disposition seems in no way unusual: many speakers, even those with significant mathematical knowledge, would do the same. Similar examples can easily be generated for a wide range of mathematical vocabulary.

These cases are suggestive, but the proponent of the 2D account of the cognitive significance phenomena might reasonably object that they do not directly bear on her view. She might point out that her account does not require there to be large numbers of apriori representations; perhaps there are only a few cognitive trivialities. But cases can be developed that undermine the apriority even of “Hesperus = Hesperus”:

**Identity Claims** Perhaps competent speakers use the identity sign in such a way that it can be applied correctly only to middle-sized dry goods, and not to objects as large as planets. Surely it is conceivable that some of us would defer to some authority in this matter. (At least a few speakers have accepted the testimony of experts who claim that certain entities described by quantum mechanics are not self-identical, despite not clearly understanding the relevant physics themselves.)

That a thinker might accept expert testimony about what is and what is not self-identical shows that it is possible that there be a thinker for whom it is not apriori that Hesperus = Hesperus. Of course, this is a problem for the 2D account of cognitive significance only if such a thinker might continue to distinguish the thought that Hesperus = Hesperus from the thought that Hesperus = Phosphorus with respect to cognitive significance. But it seems clear that this is possible. Such a thinker could continue to be disposed to take it for granted that Hesperus = Hesperus, to find this claim obvious, to appeal to it in reasoning, and so forth, while finding the claim that Hesperus = Phosphorus surprising, and not something that can be taken for granted; and such a thinker could be rational in so doing. But if this is the case, then the Cognitive Significance Principle cannot be an analysis of cognitive significance.

Can primary intensions be used to formulate a different account of distinctions in cognitive significance for a Quinean thinker? It is true that the thinker that we have imagined

---

<sup>14</sup>This case illustrates that the cases do not depend on actual ignorance or error. Even in many cases in which we in fact have correct beliefs, it is possible that we are mistaken; and if we were corrected by experts, we would change our beliefs. This is all that is needed to get the argument off the ground.



would reject the thought that Hesperus = Hesperus in different scenarios than the thought that Hesperus = Phosphorus. But it is not clear how this fact can be used to establish that the thought that Hesperus = Hesperus is cognitively less significant. It might be claimed that there are *more* scenarios in which Hesperus = Phosphorus is falsified, or that these scenarios are more likely or more salient. But these claims are very difficult to evaluate, and in the absence of further argument, by no means plausible. There are, no doubt, very many ways that the astronomical facts could turn out to be different, but the same can be said of the facts about linguistic practice in our community; it seems likely that the set of worlds in which linguistic conventions are relevantly different has the same infinite cardinality as the set of worlds in which the astronomical facts are relevantly different. Moreover, many of us are less confident in our command of language than in our command of astronomy, so there is no obvious sense in which the discovery that we are wrong about linguistic conventions would be less likely or more surprising than the discovery that we are wrong about astronomical fact.

So far, I have focused on the account of cognitive significance and cognitive triviality. But the same style of objection can be generalised to the account of rationally transitions in thought, using other cases:

**Logical truths** Timothy Williamson (2003; 2006; 2007) has described a number of cases in which speakers use logical vocabulary in a slightly deviant way. Williamson emphasises the case of Vann McGee (1985), who holds on the basis of plausible-seeming counterexamples that ‘If..then’ in English does not support modus ponens. He rejects modus ponens on these grounds. I do not know whether McGee’s claims about English are correct; and perhaps if McGee were ideally rational and acquainted with all of the empirical facts, he would give up his views. But I can easily imagine a language much like English for which McGee’s semantics is correct. One can imagine a thinker who, when she is presented with evidence (for example, the testimony of many speakers) that shows that her own language is like McGee supposes English to be, and would come to reject some instances of modus ponens. Given other patterns of usage, she would come to reject other instances of modus ponens, so for each instance, there is a scenario that would convince her to reject it.

Consider a deductive inference from ‘If Hesperus is a star, then Hesperus is bright’ and ‘Hesperus is a star’ to ‘Hesperus is bright’. Given the second part of the Cognitive Significance Principle, such an inference is rationally conclusive only if every world in which the primary intensions of the premises are true is world in which the primary intension of the conclusion is true. But there are possible thinkers who are disposed to accept that ‘If...then’ does

not support modus ponens when presented with certain empirical evidence about linguistic usage. These thinkers will not accept that every world in which the primary intensions of the premises are true is a world in which the conclusion is true; they will maintain that worlds in which modus ponens fails for this sort of conditional claim are counterexamples.

The Cognitive Significance Principle will predict that such thinkers perform no rationally conclusive inferences. The proponent of 2D might claim that this is simply the right result. But such thinkers might still display the sort of inferential behaviour that motivated the account in the first place; for example, they might be disposed (like Hammurabi) to infer from “Hesperus is bright” and “Hesperus appears in the evening sky” to “Hesperus is a bright object that appears in the evening sky”, but not from “Phosphorus is bright” and “Hesperus appears in the evening sky” to the same conclusion, and these dispositions might be rational. In the original Game Show case, the 2D explanation of the contrast between these two inferences turned on the claim that the first but not the second is rationally conclusive. Can we use the resources of primary intensions to give an alternative account of the contrast for the Quinean thinker? As before, it might be claimed that there are fewer scenarios in which linguistic usage is relevantly different, or that these scenarios are less likely. But again, though these claims are difficult to evaluate, there seems to be little reason to think that they are correct. Alternatively, it might be claimed that the sort of scenarios described so far (involving change accepting testimony from experts on matters that reflect community linguistic practice) are for some reason irrelevant to cognitive significance and to rational transitions in thought.<sup>15</sup> But Burge (1986) argues that there are closely related cases that do not essentially involve testimonial evidence from experts. And we can devise cases that bear directly on the issues at hand:

**Token Separation** I believe that Barry Smith is an ontologist who works in Buffalo, and also that Barry Smith is a philosopher of mind who works in London. At one point, I was prepared to infer from these premises that someone is an ontologist who works in London. But learned that it is unlikely that any one person could develop such a distinguished career in two distinct fields in this way. I judged that Barry Smith is not Barry Smith, and that the inference I used to accept is invalid.

At one point, I regarded the described inference as obviously valid, and hence to treat it like Hammurabi’s Inference A (and not at all like (say) an inference from the claim that Barry Smith is an ontologist who works in Buffalo and the claim that the Director of the Institute of Philosophy is a philosopher of mind who works in London). But I am disposed to give

---

<sup>15</sup>I return to a related proposal in section 2.4 below.

up the idea that inference is valid when presented with certain empirical evidence about human capacities to succeed in multiple fields. It is not clear how primary intensions can be used to give an account of this. And, of course, a person might have similar dispositions to give up identity claims that seemed obvious and trivial. (For example, some might find it conceivable that their two token uses of a name in an identity claim have different causal origins, so that one picks out the London Barry Smith and the other picks out the Buffalo Barry Smith.)

Of course, these considerations do not show that a notion of ‘rationally conclusive inference’ cannot be *stipulatively defined* in the way suggested by the Cognitive Significance Principle. In that sense, it does not show that the 2D account is *false*. But if ‘rationally conclusive inference’ is intended to be a category that does useful work in an account of our actual inferential behaviour, the described cases are extremely problematic for it. This is the theme I return to in section 2.3 below.

There are at least three ways that a proponent of the 2D account construed as a metaphysically necessary account might respond. She might claim that the Quinean dispositions described are *irrational*, so that they would not persist under rational idealisation and so would not impact on primary intension. She might try to argue that speakers with the dispositions described might have (or at least be in a position to form) multiple concepts of (say) identity or the conditional, and that some of these concepts might put them in a position to formulate thoughts with necessary primary intensions. Or she might revise the account of the cognitive significance phenomena so as to rule out the cases in question. I address the first two objections here, and return to the third in section 2.4 below.

### **Are the Dispositions Irrational?**

Is it rational to revise one’s beliefs in the face of testimony from experts, in the way many of these cases would require? One could certainly make case that it is not rationally *required* to revise ones beliefs. For example, if Bert chooses to stick to his idiosyncratic use of ‘arthritis’ in the face of expert criticism, it is at least not obvious that he is making an epistemic mistake. (In the long run, there may be negative epistemic consequences to his choice, to the extent that he is less able to communicate with other speakers. But even if this is irrational in some sense, this does not seem to be the sort of irrationality at issue.) But what the two-dimensionalist needs to show is that such revision is not rationally *permissible*.

There are some cases in which this is plausible: for example, it is hard to imagine a scenario in which it is rational for me to revise my first-order non-metalinguistic beliefs in the face of seeming experts who insist that ‘consciousness’ refers not to a mental phenomenon

but to a sort of sandwich, or that ‘twelve’ is exclusively a name for a popular hairstyle. (In any case, the Burge arguments would not apply in these cases; it is natural neither to attribute (non-metalinguistic) disagreement between me and someone who held the views in question, nor would I seem rational if I did change my non-metalinguistic opinions.) But the cases described above are not like this. It is hard to see any intuitive motivation for the claim that the subjects described are irrational. As I have already acknowledged, there are principled arguments that purport to show that certain dispositions to reject claims in the face of empirical evidence would be irrational. But it is hard to see how the proponent of 2D could accept any argument of this sort with respect to the Burge scenarios. It is simply a consequence of Chalmers’s way of generating primary intensions that subjects who are disposed to accept expert testimony in the ways described have representations with non-necessary primary intensions. Since 2D therefore entails that Quinean thinkers are possible, if it could be shown that a Quinean thinker is irrational, this would actually be a problem for 2D. (I return in section 2.4 to the question of whether it is possible to modify the account of primary intensions to avoid this consequence.)

### **Could We Have Multiple Concepts?**

Could it be that those who are willing to defer have *multiple* conditional concepts (or multiple grasps of a conditional concept), one of which is such that instances of modus ponens involving it are apriori valid? This is possible. But this saves the 2D view only if *every* rational inferer who is willing to defer has two such concepts, and only if these thinkers *only ever* deductively infer using the conditional according to which modus ponens is apriori valid. I do not see how these claims could be defended. Similarly, it might be possible for our massively deferential Quinean thinker to *form* a concept of which she can know apriori by 2D lights that it is the material conditional, or some other modus ponens supporting conditional. Perhaps she could introduce such a concept by stipulation. But it is also possible to imagine our deferential Quinean thinker as having made no such stipulations, and hence as having no non-deferential terms or concepts. In this case, no sentence phrased in vocabulary she actually understands is apriori for her, and similarly, no thought she can think without grasping new concepts is apriori for her. This possibility is all the argument requires.

### **2.3 2D is not an Explanation**

I have now argued that Quinean thinkers are possible, and that this shows that the 2D account cannot be an analysis or metaphysically necessary account of cognitive significance. I now want to argue that 2D cannot provide an explanation of cognitive significance either.

It is an empirical question what primary intensions a given thinker assigns to various representations, and so, on the assumption that Quinean thinkers are possible, it is an empirical question whether we are Quinean thinkers. My argument proceeds in two stages: I first present the case that we are Quinean thinkers. I think that this case is strong but not decisive; if it succeeds, then 2D is clearly not a successful explanation of our reasoning. I then go on to argue that even if we are not, the changes required to make us Quinean thinkers would be very small, so that if we were Quinean thinkers, our patterns of thought and reasoning would be essentially the same. This suggests that (2EC) is false, and thus that 2D is not a good explanation of the cognitive significance of our thoughts.

### 2.3.1 Are We Quinean Thinkers?

There are at least two reasons<sup>16</sup> to think that social phenomena infect all of our actual ordinary thought and discourse. The first is inductive. Consider the list of examples enumerated in the previous section. Burge-style examples can be generated even for those beliefs that have been held up as the best possible candidates for apriority. It only takes a little creativity to generate an example for any given allegedly apriori belief.

The two-dimensionalist will probably find this reasoning unpersuasive. As I noted in the previous section, it is possible for thinkers to have multiple concepts, some of which are sensitive to expert testimony and others of which are not. The two-dimensionalist might claim that we are such thinkers, so that (e.g.) we typically believe that one is not prime in two ways: one on which we would give the belief up on the basis of expert testimony, and one on which we would not. As long as we use our non-deferential concepts in deductive inference, the 2D account would survive. I personally find no introspective evidence for this multiplicity of concepts.<sup>17</sup> But this is not decisive, and the result is a stalemate.

There is, however, a second reason to think that we are actually willing to defer in a very wide range of cases. This is the centrality of testimony to our ordinary epistemic practices. Much of the recent literature on the epistemology of testimony stresses its ubiquity:

Testimony is an invaluable source of knowledge. We rely on the reports of those around us for everything from the ingredients in our food and medicine to the identity of our family members, from the history of our civilisation to the limits and contents of our planet. If we refrained from accepting what others told us, our lives, both practically and intellectually, would be unrecognisable. (Lackey,

---

<sup>16</sup>I set aside controversial (e.g. Davidsonian or Wittgensteinian) views about the nature of meaning and content that would provide further reasons if true.

<sup>17</sup>Indeed, when I consider my own case, it seems to me very likely that I am a Quinean thinker. But of course I do not expect the two-dimensionalist to be impressed by my introspection.

2008, p. 1)

Given the fact that testimony plays such a central role with respect to so many of our most crucial and closely held beliefs, it would be surprising if there were large swathes of belief that were immune to correction from testimony. And, as I have stressed, testimony is what speakers are responding to in the Burge cases.

Much testimony will not be useful to idealised reasoners considering complete descriptions of entire worlds. For example, such reasoners would not need testimony about ordinary empirical matters, since (we are supposing) they could infer the empirical truths from the description of the scenario. And certain sorts of expert testimony about mathematics and logic will be irrelevant as well (see [Chalmers \(2013, ch. 5\)](#)). For example, Philip Kitcher (1984) argues that mathematical beliefs are not apriori because they are defeasible on the basis of testimony from expert mathematicians – for example, testimony to the effect that a certain sort of a apparently valid reasoning incorporates a subtle error. But from the present perspective, this sort of testimony is irrelevant; a rationally idealised agent would be able to evaluate mathematical reasoning herself, and so would have no need to rely on the claims of mathematicians.<sup>18</sup>

Chalmers’s method of describing primary intensions appeals to an individual’s rationally idealised dispositions to apply a word or concept in a given scenario. Modulo whatever is required by rational idealisation, a particular thinker’s dispositions will be determined by contingent facts about her psychology. But given that humans rely on testimony constantly and almost automatically, it is very plausible that human psychologies are inherently disposed to accept testimony. Indeed, testimony is often preferred over other sources of evidence such as perception. (Consider how many of us accept the surprising, seemingly impossible claims of scientists about e.g. colour or solidity without really understanding the evidence on which these claims are based.) And there is a good deal of empirical evidence that willingness to accept testimony is deeply rooted in human psychology. For example, Jaswal and colleagues have shown that young children tend to trust testimony even when it contradicts their own experience ([Jaswal, 2010](#); [Jaswal et al., 2010](#)). (In related experiments, similar effects seem to vary with the age of the children; see [Clement et al. \(2004\)](#).) Gilbert and colleagues have

---

<sup>18</sup>In fact, this sort of consideration raises some subtle issues. Could there be types of mathematical reasoning that some possible thinker could complete, but that we could not (even given idealisation) without undergoing massive changes to our psychologies? For example, suppose that our reasoning capacities are limited to the Turing machine computable. Even if we were idealised by giving us unlimited time, memory, and so forth, there would still be limits to the sort of reasoning that we could perform, limits that could be overcome only by fundamentally changing our psychological natures. But there might be thinkers with computationally different psychological architecture (for examples, see [Copeland and Sylvan \(1999\)](#)) who can perform reasoning that we could not. The testimony of such thinkers might be valuable even to an agent who was as ideally rational as someone psychologically like us could possibly be.

shown that subjects under cognitive load are much more likely to misremember a sentence labeled ‘false’ as true than a sentence labeled ‘true’ as false (Gilbert et al., 1990, 1993). Gilbert argues on the basis of this and similar experiments that acceptance of testimony is a default attitude, while rejection takes work. Similar effects have been found in research on advertising. For example, Ross (1984) and colleagues have shown that celebrity testimony in advertising is effective in increasing children’s preferences for the advertised product, and that this effect is as strong in older children (who are presumably more aware of the potential ulterior motives of celebrity endorsers) as in younger. They conclude that ‘[c]ontrary to the speculation of many researchers, understanding about advertising intent and techniques and cynicism about ads had almost no influence on product preference after viewing’ (1984, p. 185). Some effects of testimony seem particularly relevant to the case where a particular view is generally held in the community. For example, it is well known that hearing a particular claim repeated tends to increase confidence in that claim (Bacon, 1979). And some researchers have argued that there are good grounds to think that humans have an evolutionarily inbuilt tendency to accept beliefs that are prevalent in their culture (Henrich and Boyd, 1998).

Of course, the subjects of some of these experiments made simple mistakes that a perfect reasoner would not make. And it is beyond doubt that we have mechanisms for detecting deception and inaccuracy of various kinds that lead us to reject testimony under certain circumstances (though it is worth noting in this context that Sperber and colleagues, in a survey of literature on the detection of misinformation, conclude, “the picture we have sketched of epistemic vigilance on a population scale is somewhat grim. Mechanisms for epistemic vigilance are not geared to filtering information transmitted on such a large scale” (2010, p. 382)). But there is good reason to think that willingness to accept testimony is a deep-rooted tendency in human psychology. This tendency would remain even when its more problematic manifestations are ironed out by idealisation. When we evaluate the way we actually use our words and concepts – their actual epistemic properties – we must take into account these facts about our psychology, and these facts suggest that we are wired to accept testimony unless we have some strong reason not to. While this is by no means a proof that we are in principle willing to accept testimonial evidence in every case, it is suggestive. In particular, to the extent that the 2D account of inference requires us to be unwilling to accept certain sorts of testimonial evidence (for example, to the effect that *modus ponens* is invalid), the two-dimensionalist owes us some empirical reason to think that our general tendencies are overcome in these cases.<sup>19</sup>

---

<sup>19</sup>There are certain special cases where it is controversial whether it is acceptable to believe on testimonial

### 2.3.2 We are Like Quinean Thinkers

The empirical case that we – or at least many of us – are Quinean thinkers is reasonably strong, but it is not decisive. Whether we are Quinean thinkers depends on subtle questions about our dispositions to respond to empirical evidence, and in particular to respond to testimony. But recall (2EC): if 2D is a (part of a) good explanation of cognitive significance (or of rational transitions in thought), then it must be the case that if we were Quinean thinkers, we would not distinguish cognitively significant and cognitively trivial thoughts (or make rational transitions in thought) in the way we do, and relatedly, that it is not possible that there be Quinean thinkers who are otherwise very similar to us and distinguish cognitively significant and cognitively trivial thoughts (or make rational transitions in thought) in the way we do. If not, then 2D is missing the truly explanatory factors; it is therefore a failed explanation, in something like the way that the “igneous rock” explanation from section 1.3 is a failed explanation.

Now consider what we would be like if we were Quinean thinkers. As the last section shows, this would require (at most) only a small and subtle change in our dispositions – so small and subtle that it is difficult to tell whether we have the required dispositions already! This small and subtle change would not change our underlying psychological nature (for example, if our thinking is a matter of computational manipulation of symbols in a language of thought, it would still be so if we were Quinean thinkers). More importantly, this small and subtle change would not make a significant difference to our dispositions to distinguish cognitively significant and cognitively trivial thoughts (or make rational transitions in thought). I therefore conclude that (2EC) is false, and hence that 2D fails as an explanation of our reasoning.

Perhaps it will be urged that the counterfactual test is unreliable; this may be so, but nonetheless it seems clear that a good explanation of our own reasoning would also cover the reasoning of Quinean thinkers. After all, many of us may well be Quinean thinkers; for example, I believe that I am. Even supposing that some of us are not, it is not plausible that the cognitive significance of their thoughts has a fundamentally different explanation than the cognitive significance of my thoughts. But this claim might be resisted on the grounds that the inferences of a Quinean thinker would be unlike the inferences of a non-

---

grounds: for example, some philosophers maintain that moral and aesthetic testimony are problematic in various respects. But the respects in which these sorts of testimony are problematic seem orthogonal to our concerns. For example, it seems problematic for me to believe testimony to the effect that Beethoven’s piano sonatas are not beautiful if I have not heard them and the testifier’s grounds are that she does not find them aesthetically appealing. But it is not obvious that there is a similar problem if the testifier’s grounds are based on (e.g.) the idea that it is a matter of conceptual or definitional truth, entities of that kind cannot be beautiful.



Quinean *rationality*, so that (despite similarities of underlying psychological organisation) the cognitive significance of their thoughts is different in kind from the cognitive significance of our thoughts, and the rationality of their inferences has a different explanation than the rationality of our inferences. In particular, it might be claimed that the Cognitive Significance Principle provides a good account of *rationality indefeasible* cognitive triviality, and of *rationality conclusive* inference. On this view, Quinean thinkers may exhibit cognitive significance of a certain sort, and may make a kind rational inferences; but these phenomena are of a different rational type than the superficially similar phenomena in us.

I have no objection to the idea that one could stipulatively define “rationally indefeasible cognitive triviality” and “rationally conclusive inference” by means of primary intensions and the Cognitive Significance Principle. But – like the notion of primary intension itself – such stipulatively defined notions are interesting only if they do some useful work. My objection, then, is that these defined notions do not offer a good explanation of the phenomena with which we began. Our original Fregean datum was that there is a rational difference between “Hesperus = Hesperus” and “Hesperus = Phosphorus”; the former seems obvious, while the latter seems surprising and in need of empirical justification. But this difference would persist whether or not “Hesperus = Hesperus” is treated as empirically indefeasible, and hence whether or not “Hesperus = Hesperus” has a necessary primary intension. (For example, suppose again that I am a Quinean thinker. I still might rationally take “Hesperus = Hesperus” for granted, find it obvious and trivial, and so forth.) This suggests that having a necessary primary intension is not what explains the distinction. Similarly, our original datum about inference is that there is a difference between Inference A and Inference B; the former seems rational (for typical thinkers) in the absence of further premises, and hence is an inference that rational thinkers would perform, while the latter does not. But again, these differences would persist (for me) whether or not I am a Quinean thinker: Inference A would continue to be rational, and would continue to be an inference that I would perform, while Inference B would not. Again, this strongly suggests that the fact that Inference A is “rationally conclusive” cannot be what explains why I make it, or why it is rational for me to do so.

## 2.4 2D Modified?

I have now argued that the 2D account of cognitive significance is false if construed as an analysis, and is a bad explanation. One possible response would be to change the account of apriority by stipulating that cases involving linguistic misunderstanding are irrelevant. (Something like this strategy is discussed in [Chalmers 2013](#), ch. 7.) For example, let LANG

be the proposition that my beliefs about linguistic usage in my community are true (i.e., the set of worlds in which linguistic usage in my community is as I actually believe it to be). For example, it could be claimed that a proposition is *quasi-apriori* for me just in case its primary intension is true in all the worlds in LANG, and that the Cognitive Significance Principle should be reformulated in terms of quasi-apriority (rather than apriority). This would alleviate many of the problems posed by the Burge cases: for example, it will plausibly turn out to be quasi-apriori for me that all bachelors are men and that one is not prime. But there are a number of reasons to resist this proposal.<sup>20</sup>

1. The clearest examples that motivate the idea that there could be Quinean thinkers involve willingness to revise ones beliefs in response to testimony. The two-dimensionalist strategy was to appeal to the actual epistemic properties of our words. But many of our words are actually epistemically responsive to correction on the basis of testimony and community usage. There is no clear justification for ignoring these epistemic facts. Attempting to stipulate the significance of testimony away is therefore ad hoc. What motivates the idea that testimony is uniquely irrelevant to the cognitive significance of my thoughts? Why can we just ignore worlds in which linguistic meaning is different?
2. The strategy would generate spurious quasi-apriorities. For example, it would turn out to be quasi-apriori (and hence cognitively trivial) for me that (for example) my linguistic community standardly uses ‘arthritis’ to pick out a certain disease. But surely this is a significant empirical matter.
3. Some falsehoods would turn out to be quasi-apriori on this view. For example, it would turn out to be quasi-apriori for Bert that arthritis can occur in the thigh, since this is so in all the worlds in which his beliefs about linguistic meaning are true. (Perhaps the view could be adapted to this consequence; after all, many views of apriori justification allow that such justification is defeasible, and hence allow for apriori justified falsehoods. But it would be a serious concession for Chalmers, who stresses the conclusive character of apriori justification (e.g., [Chalmers \(2006, p. 99\)](#)).

This modification of the 2D account fails to save the 2D account of cognitive significance. (I return to it in my discussion of conceivability, below.) But could there be some other

---

<sup>20</sup>A closely related strategy for defending 2D would define quasi-apriority\* by appealing to the primary intension that an expression would have in a nearby world where the subject used it non-deferentially. This would evade the second objection as stated, but would still be subject to the first and third objections. And it would threaten to generate other problematic quasi-apriorities\*. For example, it would turn out to be quasi-apriori\* for me on such a view that I do not defer to my linguistic community about whether arthritis can occur in the thigh. But this is both false, and an empirical matter.

modification that does the relevant work? The features of the 2D account that are targeted by the Quinean argument are quite central: (i) the idea that cognitive significance is to be analysed in terms of apriority, and (ii) that apriori justification is empirically indefeasible. It is hard to see how these features could be dropped: (i) is the very core of the account, and (ii) is guaranteed by the account of primary intensions and the Core Thesis. But the Quinean objections could be raised to any account with these features. I therefore conclude that the prospects for a view of this sort that evades the objections are not good.

### 3 Conceivability

I have so far focused on the 2D attempt to analyse or explain cognitive significance, and have argued that this attempt fails. But much of the interest in 2D in the literature focuses on the link it promises to *modality*. A plausible view has it that conceivability is evidence of possibility. But the Kripke/Putnam examples of the necessary aposteriori seem to suggest that this evidence is highly fallible. For example, suppose (with Kripke (1980) and Putnam (1975)) that it is necessary that water is identical to  $H_2O$ . It nonetheless seems conceivable that there is water without hydrogen; certainly it is hard to see how the possibility that water lack hydrogen could be ruled out on the basis of purely apriori reasoning.<sup>21</sup> So if Kripke and Putnam are correct, conceivability cannot be a good guide to possibility in many cases. And this appears to open up the possibility that aposteriori necessities can be wielded by proponents of particular putative necessary truths: for example, physicalists could claim that it is aposteriori but necessary that pain = c-fiber firing, and that the apparent conceivability of zombies or disembodied minds is misleading in the same way that the apparent conceivability of water without hydrogen is misleading.

Two-dimensionalism promises a diagnosis, and a way to tell when conceivability cannot lead one astray. Both are modeled on Kripke's own account of modal error. According to Kripke, the sort of conceivable world that might lead one to believe that there could be water without hydrogen is not a world in which there is in fact water without hydrogen – there are no such worlds – but rather a world in which some liquid that does not contain hydrogen plays the role in our lives that water actually plays, and in particular, a world in which this liquid causes the same sorts of sensations that water actually causes. Kripke goes on to claim that there are no analogously misleading worlds with respect to pain, since any phenomenon that plays the role in our lives that pain actually plays, including feeling like pain, just *is* pain (1980, e.g. pp. 148-153); so, Kripke claimed, the physicalist attempt

---

<sup>21</sup>At least, it is hard to see as long as we set aside Millian views (Salmon, 1986) according to which one can know apriori that water is  $H_2O$  by knowing that water is water.

to make use of aposteriori necessities fails.

Kripke suggests that something like this strategy is applicable in a wide range of cases, but he does not claim that it is the only possible source of illusions of possibility. (“That the usual moves and analogies are not available to solve the problems of the identity theorist is, of course, no proof that no moves are available” (Kripke, 1980, p. 155).) In general, his strategy is piecemeal: he attempts to tell a psychologically plausible story about what we were imagining in particular cases (see Yablo (2006)). The two-dimensionalist generalises and systematises Kripke’s story in the following way. We can conceive of a world in which some substance other than H<sub>2</sub>O plays the role in our lives that H<sub>2</sub>O actually plays, and in particular, a world in which this substance causes the same sorts of sensations that water actually causes. When we consider of this world as actual, we judge that our word ‘water’ (given its actual epistemic properties) applies to this substance. That is to say, the primary intension of ‘water’ in this world includes a substance that lacks hydrogen. So the primary intension of ‘Water lacks hydrogen’ is possibly true. But it is a mistake to conclude from this that it is *metaphysically* possible that water lacks hydrogen. Metaphysical possibility is a matter of *secondary* intension being possibly true. Since the primary and secondary intensions of ‘Water lacks hydrogen’ are different, one cannot infer that the secondary intension is possibly true from the fact that the primary intension is possibly true.

But, the two-dimensionalist story continues, some representations are not like this. For example, both the primary intension and the secondary intension of ‘pain’ include in each world all and only those events that feel a certain way. Since the primary and secondary intensions are the same, one may infer from the possibility of the primary intension to metaphysical possibility. So the physicalist appeal to aposteriori necessities fails. On this view, then:

**Conceivability Principle** A representation is conceivably true just in case its primary intension is possibly true.

**Possibility Principle** A representation is metaphysically possible just in case its secondary intension is possibly true.

**Conceivability-Possibility Link** The conceivability of a representation is a reliable guide to its metaphysical possibility if and only if its primary intension is identical to its secondary intension.

It is important to be clear about what the account does and does not explain. If correct, it would explain *why illusions of conceivability occur*, and *when conceivability evidence can*

*be relied upon.* It would *not* explain *why conceivability is evidence of possibility of primary intension*; this is something that the account takes for granted. It therefore does not explain *why conceivability bears on possibility at all.*

Is the account successful? I will argue that the Quinean considerations from the previous section show that the matter is not so simple. Consider a thinker – perhaps yourself – who is willing to defer to other speakers on some paradigm apriori matter – say, whether one is prime. Since you are willing to defer, the primary intension that you associate with “one is prime” is possibly true, so by the Conceivability Principle, it is conceivable (for you) that one is prime. Now, either the primary intension that you associate with “one is prime” is the same as its secondary intension, or it is not. Suppose it is. Then, by the Conceivability-Possibility Link, it is metaphysically possible that one is prime. But this is an unacceptable result; mathematical truths are necessary if anything is. So the primary intension of “one is prime” must be distinct from its secondary intension.

Now this line of reasoning can be generalised in any case in which a person has differential dispositions. But many or most of us have such dispositions with respect to the concepts we use to think about our experiences. For example, some people who are under the misimpression that nausea is a kind of pain. No amount of introspective or neurobiological evidence will convince them otherwise, but testimony from other expert speakers will. Conversely, most of us would probably bow to the expertise of a large number of expert speakers (armed with dictionaries and the like) if they were to claim that nausea is a type of pain after all. Similarly, I take it that the sensation associated with hearing a very loud noise is a kind of pain. But it is conceivable that other expert speakers do not take this sensation to be a kind of pain, and if they did, I would defer to them. Now suppose that I feel a pain in my ears from having heard a very loud noise. In that case, the primary intension of “There is a physical duplicate of me who is not pain” will be possible for me, since it will be true in a world considered as actual in which other speakers correct my usage and I accept their correction. So I am in pain, but it is conceivable that a physical duplicate of me is not in pain. As in the “one is prime” case, there is a choice: either we can claim that the primary intension of “There is a physical duplicate of me who is not in pain” is the same as its secondary intension, or we can deny this. And as before, I take it that denying this is the more reasonable option.

I have already argued that many of us have dispositions to accept testimony of this kind in a wide variety of cases. In all of the cases where we have this sort of disposition, this style of argument will show that primary and secondary intensions come apart. But then if the Conceivability-Possibility Link is correct, in all of the cases where we have such dispositions,

conceivability is not a reliable guide to possibility. The more we are like Quinean thinkers – and as I have argued, we are very like them – the less conceivability would reveal about possibility. And in particular, since we (or at least most of us) have deferential dispositions with respect to concepts like “pain”, conceivability would be no evidence of the possibility of zombies.

Now, I don’t take this to show that conceivability is no evidence for the possibility of zombies. It is pretty clear that this is not the sort of scenario that we conceive of when we judge that zombies are possible. So the argument does *not* show that conceivability is in fact unreliable in this case. But it *does* show that the 2D account has problematic consequences – at least if it is applied in a straightforward way. The 2D account predicts that conceivability is no evidence of possibility in many cases *for the wrong reasons*. It may or may not be the case that conceivability is good evidence of the possibility of zombies, but even if it is not, this is not because of our dispositions to accept testimony.

Can the account be fixed? A first idea would be to systematically rule out worlds in which linguistic meaning changes, for example in something like the way mooted in section 2.4. On this view, a representation would be conceivable for a thinker just in case its primary intension is true in some world in which the thinker’s beliefs about linguistic practices in his community are true. But this cannot save the 2D account of conceivability for the same sorts of reasons that it cannot save the 2D account of cognitive significance. As before, in the absence of further motivation, the restriction seems ad hoc. More seriously, the account would rule inconceivable much that is in fact conceivable. For example, it seems clear that conceivability is good evidence for the claim that linguistic practices in our community could have been different; but on the proposed account, this would not be conceivable.

Given that we have a good sense of when conceivable scenarios are and are not relevant to possibility, can we just disregard them as appropriate? For example, can we simply ignore the case where other speakers do not apply “pain” to sensations caused by very loud noise when we are considering the possibility of zombies, but not when we are considering the possibility of linguistic change? Of course, we can do this. But this is not a result delivered by 2D; we have at this point given up the idea that 2D provides a general recipe for deciding which scenarios are and aren’t relevant to a given possibility claim. We are now back to Kripke’s piecemeal method. And this undermines 2D’s claim to have illuminated the relationship between conceivability and possibility. We no longer have a systematic explanation of when and why illusions of possibility occur: there are cases in which primary and secondary intension come apart in which there is no serious illusion of possibility. And we no longer have an explanation of why conceivability evidence is reliable when it is: the

account predicts that conceivability will be unreliable with respect to representations about which we are disposed to accept testimony, but this does not seem to be the case.

Thus 2D does not explain the relationship between conceivability and possibility. Conceivability may be perfectly good evidence of possibility for all that has been shown; but this fact is not illuminated by appeal to primary intension.

## 4 Conclusion

For any particular putatively apriori truth  $p$ , the two-dimensionalist can respond to a Quinean or Burgean argument that  $p$  would be rejected (by us, or by some particular thinker) under certain circumstances by accepting the claim that  $p$  is not apriori (for us, or for that thinker) as a discovery. But the Burgean considerations threaten to generalize: it is possible some thinker (in most respects much like us) has deferential dispositions with respect to all of her concepts, and for her, nothing would be apriori. There is even reason to take seriously the possibility that we are such thinkers. But this shows that primary intensions have no useful role to play in an analysis or explanation of our reasoning, and provide no insight into the relation between conceivability and possibility. Nor could any successful account be constructed from the same materials – i.e., from an indefeasible notion of apriority.

Perhaps 2D can be shown to have some other application – perhaps as a ‘picture’ of our thinking and reasoning on the model of Kripke’s (1980) causal picture of reference, or as an idealised model. But until the applications and limitations of such are made clear, there is no reason to think that primary intensions are an interesting form of narrow content.

## References

- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5:241–252.
- Block, N. and Stalnaker, R. (1999). Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review*, 108:1–46.
- Burge, T. (1979). Individualism and the mental. In Ludlow, P. and Martin, N., editors, *Individualism and Self-Knowledge*, pages 21–84. CSLI, Stanford. Originally published in *Midwest Studies in Philosophy*, 4, 1979.
- Burge, T. (1986). Intellectual norms and the foundations of mind. *Journal of Philosophy*, 83:697–720.

- Burge, T. (1993). Content preservation. *The Philosophical Review*, 102:457–488.
- Byrne, A. and Pryor, J. (2006). Bad intensions. In [García-Carpintero and Macià \(2006\)](#).
- Carnap, R. (1955). Meaning and synonymy in natural languages. *Philosophical Studies*, 6:33–48.
- Carnap, R. (1956). *Meaning and Necessity, second edition*. University of Chicago Press, Chicago.
- Chalmers, D. (1996). *The Conscious Mind*. Oxford University Press, New York.
- Chalmers, D. (2002a). The components of content. In Chalmers, D., editor, *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press, New York.
- Chalmers, D. (2002b). On sense and intension. *Philosophical Perspectives*, 16.
- Chalmers, D. (2006). The foundations of two-dimensional semantics. In [García-Carpintero and Macià \(2006\)](#).
- Chalmers, D. (2007). Two-dimensional semantics. In Lepore, E. and Smith, B., editors, *Oxford Handbook of the Philosophy of Language*. Oxford University Press, Oxford.
- Chalmers, D. (2011a). Propositions and attitudes: A Fregean account. *Nous*, 45:595–639.
- Chalmers, D. (2011b). Revisability and conceptual change. *The Journal of Philosophy*, 108.
- Chalmers, D. (2013). *Constructing the World*. Oxford University Press, Oxford.
- Clement, F., Koenig, M., and Harris, P. (2004). The ontogenesis of trust. *Mind and Language*, 19:360–379.
- Copeland, B. J. and Sylvan, R. (1999). Beyond the universal turing machine. *Australasian Journal of Philosophy*, 77:46–67.
- Fodor, J. (1990). A theory of content, parts 1 and 2. In *A Theory of Content and Other Essays*. The MIT Press, Cambridge, MA.
- García-Carpintero, M. and Macià, J., editors (2006). *Two-Dimensional Semantics*. Oxford University Press, New York.
- Gilbert, D. T., Krull, D. S., and Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 69:601–613.



- Gilbert, D. T., Tafarodi, R. W., and Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, 65:221–233.
- Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell, Oxford.
- Henrich, J. and Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19:215–241.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford University Press, New York.
- Jaswal, V. K. (2010). Believing what you're told: Young children's trust in unexpected testimony about the physical world. *Cognitive Psychology*, 61:248–272.
- Jaswal, V. K., Croft, A. C., Setia, A. R., and Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, 21:1541–1547.
- Kaplan, D. (1977). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes From Kaplan*, pages 481–564. Oxford University Press, New York.
- Kitcher, P. (1984). *The Nature of Mathematical Knowledge*. Oxford University Press, New York.
- Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In *Scientific Explanation*, pages 410–505. University of Minnesota Press, Minneapolis.
- Kripke, S. A. (1980). *Naming and Necessity*. Harvard University Press, Cambridge, MA.
- Lackey, J. (2008). *Learning From Words: Testimony as a Source of Knowledge*. Oxford University Press, New York.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70:556–567.
- Macpherson, G., editor (2002). *Black's Medical Dictionary, 40th edition*. A & C Black, London.
- Mates, B. (1951). Analytic sentences. *The Philosophical Review*, 60:525–534.
- McGee, V. (1985). A counterexample to modus ponens. *Journal of Philosophy*, 82:462–471.
- Neurath, O. (1959). Protocol sentences. In Ayer, A., editor, *Logical Positivism*, pages 199–208. The Free Press, New York.

- Nozick, R. (1981). *Philosophical Explanations*. Harvard University Press, Cambridge, MA.
- Putnam, H. (1975). The meaning of 'meaning'. In *Mind, Language, and Reality: Philosophical Papers Volume 2*, pages 215–271. Cambridge University Press, New York. Originally published in *Minnesota Studies in the Philosophy of Science*, 7, 1975.
- Quine, W. V. O. (1953). Two dogmas of empiricism. In *From a Logical Point of View*, pages 20–46. Harvard University Press, Cambridge, MA.
- Ross, R. P., Campbell, T., Wright, J. C., Huston, A. C., Rice, M. L., and Turk, P. (1984). When celebrities talk, children listen: An experimental analysis of children's responses to tv ads with celebrity endorsement. *Journal of Applied Developmental Psychology*, 5:185–202.
- Salmon, N. (1986). *Frege's Puzzle*. Ridgeview, Atascadero, CA.
- Schaffer, J. (2000). Trumping preemption. *Journal of Philosophy*, 97:165–181.
- Schiffer, S. (2003). *The Things We Mean*. Oxford University Press, Oxford.
- Schroeter, L. (2003). Gruesome diagonals. *Philosophers' Imprint*, 3:1–23.
- Schroeter, L. (2004). The rationalist foundations of chalmers's 2-d semantics. *Philosophical Studies*, 118:227–255.
- Schroeter, L. (2005). Considering empty worlds as actual. *Australasian Journal of Philosophy*, 83:331–347.
- Scriven, M. (2008). *Depth*. Harvard University Press, Cambridge, MA.
- Soames, S. (2007). *Reference and Description: The Case against Two-Dimensionalism*. Princeton University Press, Princeton.
- Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., and Wilson, D. (2010). Epistemic vigilance. *Mind and Language*, 25:359–393.
- U.S. Department of Health and National Institutes of Health (2004). Handout on health: Rheumatoid arthritis.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press, Oxford.
- Williamson, T. (2003). Blind reasoning. *Aristotelian Society Supplementary Volume*, 77:249–293.

Williamson, T. (2006). Conceptual truth. *Aristotelian Society Supplementary Volume*, 80:1–41.

Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell, Oxford.

Woodward, J. and Hitchcock, C. (2003). Explanatory generalizations, part i: A counterfactual account. *Noûs*, 37(1):1–24.

Yablo, S. (2002). Coulda, woulda, shoulda. In Gendler, T. S. and Hawthorne, J., editors, *Conceivability and Possibility*, pages 441–492. Oxford University Press, New York.

Yablo, S. (2006). No fool's cold. In [García-Carpintero and Macià \(2006\)](#), pages 181–204.