

# Epistemic two-dimensionalism

PHIL 93507

Jeff Speaks

December 1, 2009

1	Four puzzles . . . . .	1
2	Epistemic two-dimensionalism . . . . .	3
	2.1 Two-dimensional semantics . . . . .	3
	2.2 Epistemic intensions . . . . .	3
	2.3 Excluded vocabulary and the construction of scenarios . . . . .	4
3	Epistemic two-dimensionalism and the four puzzles . . . . .	5
	3.1 The a priori . . . . .	5
	3.2 Cognitive significance . . . . .	10
	3.2.1 Structured epistemic intensions . . . . .	11
	3.2.2 Structured epistemic intensions + logical truth . . . . .	12
	3.2.3 Three-dimensionalism? . . . . .	12
	3.3 Substitution failures . . . . .	13
	3.3.1 A simple semantics for attitude ascriptions . . . . .	13
	3.3.2 The semantics once-complicated . . . . .	14
	3.3.3 The semantics twice-complicated . . . . .	16
	3.3.4 A fallback position . . . . .	16
	3.4 The problem of the essential indexical . . . . .	17
4	Conclusion . . . . .	19

Nowadays, most two-dimensionalists are not contextual two-dimensionalists, but hold a variant of the view that Chalmers has called *epistemic two-dimensionalism*. As with the contextual 2Dist, the epistemic 2Dist introduces a second dimension of meaning to explain the epistemic properties of sentences; as with contextual 2Dism, it is easiest to understand the point of the view by thinking about what this extra dimension of meaning is supposed to explain.

## 1 Four puzzles

A Millian-Russellian semantic theory is one according to which the meanings of proper names are the objects for which they stand, and the meanings of predicates are the properties (or relations) they express. Given a plausible compositionality principle (which I will assume), the Millian and Russellian must hold that sentences which differ only in

the substitution of proper names which have the same reference (relative to the relevant context) must express the same proposition.

A Fregean semantic theory, which identifies meanings with modes of presentation of, or ways of thinking about, Russellian contents, will make distinctions in meaning that the Millian-Russellian does not.

A good way to see the motivations for a Fregean theory of content is by noting four puzzles for Millian-Russellian theories of content. Each takes the form of an argument that a pair of sentences express different propositions; and in each case the Millian-Russellian is forced to say that they express the same proposition.

*Puzzle 1: the a priori*

The proposition expressed by [1] is knowable a priori, whereas the proposition expressed by [2] is not.

[1] If Hesperus exists, then Hesperus is Hesperus.

[2] If Hesperus exists, then Hesperus is Phosphorus.

*Puzzle 2: cognitive significance*

The proposition expressed by [1] is trivial ('lacks cognitive significance'), whereas the proposition expressed by [2] is non-trivial ('has cognitive significance'), and might be something that one discovers to be true.

*Puzzle 3: substitution failures*

It is easy to imagine cases in which the proposition expressed by [3] is true whereas the proposition expressed by [4] is false; so they must be distinct.

[3] John believes that Hesperus is visible in the evening.

[4] John believes that Phosphorus is visible in the evening.

*Puzzle 4: the problem of the essential indexical*

The following sentences can't express the same proposition, since we can imagine circumstances in which my coming to believe the proposition expressed by the first would lead to immediate action on my part, whereas my coming to believe the proposition expressed by the second would not.

[5] I am on fire.

[6] Jeff Speaks is on fire.

What would it take for the Fregean to solve these puzzles, and explain these differences between the pairs of sentences above? The standards we should apply here are the same as we should apply to any attempted explanation of a class of facts; I take it that these include the following two:

1. *Extensional adequacy.* The Fregean should provide distinctions in content corresponding to the intuitive differences above.

2. *Non-circularity*. The Fregean should not define her explanans — Fregean contents — in terms of the explananda (differences with respect to a prioricity, cognitive significance, etc.).

The non-circularity condition is intuitive but a bit hard to state clearly. The basic idea can be brought out by a (made-up) example.

Suppose one wants to explain puzzle 3 — apparent substitution failures involving coreferential names in attitude ascriptions — and does so by claiming that names have senses which are defined by the following feature: two names have a different sense iff it is possible for substitution of one for the other in the complement of an attitude ascription to change speakers' intuitions about truth-value. This is a perfectly fine thing to say about the senses of names; but if this is really what defines difference in sense, one cannot plausibly claim to have explained why speakers take there to be a difference in truth-value between sentences like [3] and [4].

## 2 Epistemic two-dimensionalism

### 2.1 *Two-dimensional semantics*

The newest, and perhaps most currently popular, forms of Fregeanism are various forms of two-dimensionalist semantics.

For our purposes, we can think of two-dimensionalist semantics as the view that a sentence has, relative to a context, (at least) two sorts of meaning, one of which — its primary intension — is more closely related to the sentence's epistemic or cognitive properties, and one of which — its secondary intension — is more closely linked to the sentence's (metaphysical) modal profile. All two-dimensionalists endorse the following two theses:

A sentence is *a priori* (i.e., epistemically necessary) iff it has a necessary primary intension.

A sentence is metaphysically necessary iff it has a necessary secondary intension.

Because secondary intension just is Millian-Russellian content, the second thesis is common ground between two-dimensionalists and their opponents. The first thesis is what David Chalmers calls the 'Core Thesis' of two-dimensionalism; different versions of two-dimensional semantics are defined by their different views of the nature of primary intensions.

### 2.2 *Epistemic intensions*

*Epistemic two-dimensionalism* is a version of two-dimensionalism which David Chalmers has defended in an important recent series of papers, and it is, I think, probably the most widely endorsed version of two-dimensionalist semantics.

According to epistemic two-dimensionalism, primary intensions are *epistemic intensions*. Epistemic intensions, like all intensions, are functions from indices of a certain sort to extensions.

The relevant indices are epistemic possibilities — or, as Chalmers says, *scenarios*. We can think of a scenario as a conjunction of sentences which is both epistemically possible — in the sense that its negation is not a priori — and epistemically complete — in the sense that for any sentence, the scenario epistemically necessitates either it or its negation.

The extension of an expression with respect to a scenario is what we would come to know its extension to be, were we to learn that this scenario is actual. As Chalmers puts it,

“there is a strong epistemic dependence of an expression’s extension on the state of the world. If we come to know that the world has a certain character, we are in a position to conclude that the expression has a certain extension. And if we were to learn that the world has a different character, we would be in a position to conclude the expression has a different extension. That is: we are in a position to come to know the extension of an expression, depending on which epistemic possibility turns out to be actual.” (Chalmers (2004), 177-8)

So a sentence is true with respect to a scenario if and only if learning that that scenario is actual would put us in a position to know that the sentence is true — if, in Chalmers’ terms, the sentence is *epistemically necessitated* by that scenario.

Chalmers discusses several ways of making the notion of epistemic necessitation precise but, given the Core Thesis, it turns out that epistemic necessitation must be equivalent to a priori consequence, in the following sense (proof at the end for anyone interested):

For any scenario  $e$ , and any sentence  $S$ ,  $e$  epistemically necessitates  $S$  iff  $\ulcorner e \rightarrow S \urcorner$  is a priori.

### 2.3 Excluded vocabulary and the construction of scenarios

We are thinking of scenarios as long interpreted sentences. But, for two reasons, certain sorts of vocabulary must be systematically excluded from the construction of scenarios:

- The problem of trivializing epistemic intensions.
- The need to preserve a link between conceivability and possibility by guaranteeing that every scenario describes a metaphysically possible world, while providing enough scenarios for the Core Thesis to come out true.

For our purposes, the most important class of excluded vocabulary will be proper names. But even though scenarios don’t include names, they will contain various definite descriptions which denote different objects. Given this, for a name  $n$  to have a reference in the scenario is for the scenario to epistemically necessitate the claim that the referent of the name is the same as the object denoted by one (or more) of these definite descriptions. In

other words, for  $n$  to have a reference in a scenario  $e$  is for there to be some description ‘the  $F$ ’ in the scenario such that the conditional  $(e \rightarrow (n = \text{the } F))$  is a priori.<sup>1</sup>

### 3 Epistemic two-dimensionalism and the four puzzles

#### 3.1 *The a priori*

Suppose, to continue with the standard case, that you learn the name ‘Hesperus’ the old fashioned way, by having it demonstrated as the brightest body in a certain part of the night sky, and ‘Phosphorus’ in the same way, but with the demonstration being to a certain part of the morning sky. Then, intuitively, the names should have different epistemic intensions for you. Taking as actual a scenario which contains sentences like ‘The brightest body in the evening sky is a planet’ you will, the idea goes, take ‘Hesperus is the brightest body in the evening sky’ to be true. And, furthermore, you might do the same with ‘Phosphorus’ and the description ‘the brightest body in the morning sky’ — even if the scenario also contains the sentence ‘The brightest body in the morning sky is not the brightest body in the evening sky.’ So, relative to such a scenario, the epistemic intensions of the two names, as used by you, will determine a different reference. So there will be at least one scenario with respect to which [2] is false, and so [2] will come out a posteriori — which is what we want.

Nonetheless, the epistemic two-dimensionalist treatment of the a priori is open to serious objections.

*Objection 1: the epistemic two-dimensionalist is forced to make even more shocking claims about the a priori than the Millian-Russellian.*

To set up the argument, let’s say that two properties,  $F$  and  $G$ , are *independent* if and only (1)  $F$ ’s being instantiated is compossible both with  $G$ ’s being instantiated and its being uninstantiated, and (2)  $F$ ’s being instantiated does not a priori entail either  $G$ ’s being instantiated, or  $G$ ’s being uninstantiated.

This definition of independence has a natural extension to the case of concrete particulars: two things,  $m$  and  $n$ , are independent iff (1)  $m$ ’s existence is compossible both with  $n$ ’s existence and with  $n$ ’s nonexistence, and (2)  $m$ ’s existence does not a priori entail either  $n$ ’s existence, or  $n$ ’s nonexistence.<sup>2</sup>

Given this, I suggest the following as two plausible claims about the limits of a priori knowledge:

- [A] On the basis of the knowledge that some particular thing  $n$  is  $F$ , you can’t know a priori whether some other thing exists which instantiates some other property  $G$ , if the two properties are independent.

<sup>1</sup>This follows the discussion in Chalmers (2006), §3.7.

<sup>2</sup>And, obviously, we can extend the definition to the case of a pair of a particular object and a property. A particular  $n$  and property  $F$  are independent iff (1)  $n$ ’s existence is compossible both with  $F$ ’s being instantiated and with  $F$ ’s uninstantiated, and (2)  $n$ ’s existence does not a priori entail either that  $F$  is instantiated, or that  $F$  is uninstantiated.

[B] Given knowledge that some particular thing  $n$  is  $F$ , you are not in a position to know a priori which of two epistemically possibly hypotheses about some other particular thing  $m$  — say, that  $m$  is  $G$  and that  $m$  is  $H$  — is true, so long as  $F$ ,  $G$ , and  $H$  are independent.

I'll now argue that epistemic two-dimensionalism is committed to denying both [A] and [B].

Consider my use of the name 'Mick Jagger.' Suppose that the properties I attribute to Mick Jagger can be divided into two conjunctions of properties,  $F$  and  $G$ :

$F$ : the lead singer on *Sympathy for the Devil* & member of the greatest band of the 1960's & once enrolled in the London School of Economics & the star of *Freejack* & ...

$G$ : member of the greatest band of the 1960's & one of the Glimmer Twins & born in Kent in 1943 & helped to organize the infamous concert at Altamont & ...

What matters is that there is a rough parity between the two as regards the number of properties in the two conjunctions and the centrality of those properties to my conception of Mick Jagger. It seems clear that  $F$  and  $G$  could be *independent*, in the sense defined at the outset:  $F$ 's being instantiated is compossible with  $G$  either being or not being instantiated, and one can't know a priori that if  $F$  is instantiated,  $G$  is, or the reverse.

As a consequence, the following is epistemically possible:

[7] Something is the  $F$ , and something else is the  $G$ .

For any scenario which epistemically necessitates [7], will there be a description 'the  $D$ ' such that 'Mick Jagger is the  $D$ ' is epistemically necessitated by the scenario? Parity reasoning indicates not. There are two candidate references for 'Mick Jagger' in the scenario; nothing will favor one over the other, and so it will not be epistemically necessary that either is the referent of the name. So, in scenarios of this sort, the epistemic intension of 'Mick Jagger' (for the relevant person) will deliver no reference. So, presumably,

Mick Jagger exists.

will be false at these scenarios. But then it looks like the material conditional

[8] Mick Jagger exists  $\rightarrow \neg$  (something is the  $F$ , and something else is the  $G$ )

will have an epistemic intension which is true at every scenario, since any scenario at which the antecedent is true will not be a scenario in which something is  $F$  and something else is  $G$ , and hence will be one in which the consequent of the conditional is true. So, by the Core Thesis, this conditional will be a priori. But this is clearly incorrect; given knowledge that Mick Jagger exists, one cannot deduce a priori that it is not the case that one thing is  $F$ , and something else is  $G$ .

The point which this example illustrates is not restricted to names; it works just as well for any sort of vocabulary which might be excluded from scenarios. The point is this: for many expressions  $n$ , there will be some sentence  $S$  such that (i) any epistemic possibility in which  $S$  is true is one in which  $n$  lacks a reference, and so in which ' $n$  exists' is false, and yet (ii) ' $n$  exists  $\rightarrow \neg S$ ' is clearly a posteriori. But (i) and (ii) are inconsistent with the conjunction of the Core Thesis and epistemic two-dimensionalism. If (i) is true, then any epistemic possibility in which ' $n$  exists' is true is one in which ' $\neg S$ ' is true, from which it follows that the material conditional ' $n$  exists  $\rightarrow \neg S$ ' is true in every scenario. So, by the Core Thesis, it is a priori — which contradicts (ii).

If the foregoing line of argument is correct, another way of making the same point is that the sentence

[9] Mick Jagger is the  $F$ , and there is someone else who is the  $G$ .

will never be true at a scenario, since there is no scenario at which 'Mick Jagger' has a reference and one thing is  $F$ , while something else is  $G$ . It follows that [9] is not epistemically possible, and hence that its negation is a priori:

[10]  $\neg (\text{Mick Jagger is the } F) \vee \neg \exists x Gx$ .

But the claim that [10] is a priori violates claim [A] about the a priori above, since if [10] is a priori, given knowledge that Mick Jagger is  $F$ , one can know a priori that nothing is  $G$ . But this is surely a straightforward case of deducing a priori the nonexistence of something with certain properties from knowledge of the properties of something else, despite the fact that the relevant properties are independent.

A different sort of puzzling case results from taking into account two names which are such that the properties I associate with the two names overlap. For example, consider my use of the name 'Keith Richards.' As it happens, we can split the properties I associate with the name into two complex properties, one of which is the property  $G$  discussed in connection with Mick Jagger above:

member of the greatest band of the 1960's & one of the Glimmer Twins &  
born in Kent in 1943 & helped to organize the infamous concert at Altamont  
& ...

while the other is a distinct property  $H$  of the same sort. As above, it might well be that  $F$ ,  $G$ , and  $H$  are independent properties. Suppose we know

[11] Keith Richards is the  $G \vee$  Keith Richards is the  $H$ .

This is presumably a posteriori, as is

[12] Mick Jagger is the  $F$ .

But, given the above argument that [10] is a priori, we can know a priori that if [12] is true, nothing which is not Mick Jagger exists which is  $G$ . But then it follows that if we know just [11] and [12] along with the proposition that Keith Richards is not Mick Jagger, we are in position to deduce a priori that Keith Richards is  $H$ . That is, the following is a priori:

[13] (Mick Jagger is the  $F$ )  $\rightarrow$  ((Keith Richards is the  $G \vee$  Keith Richards is the  $H$ )  $\rightarrow$  Keith Richards is the  $H$ )

In other words, given some knowledge about Mick Jagger's properties, we're able to rule out a priori an otherwise epistemically possible hypothesis about Keith Richards' properties — even though the relevant properties are independent. But, as above, this seems odd (as well as a violation of claim [B] about the a priori); how could knowing some stuff about Mick Jagger put me in a position to decide between two conflicting hypotheses about what Keith Richards is like, if the properties attributed to Keith by either hypothesis are, in the above sense, independent of the known properties of Mick Jagger?

The epistemic two-dimensionalist might, of course, simply deny the principles [A] and [B] about the a priori with which we began, and accept the conclusion that the sentences discussed above are a priori. She might argue as follows, focusing on the example of [10]:

Consider the claim that if Mick Jagger is  $F$ , and I stand in no relevant indexical relations to Mick Jagger, there can't be something else, which is not Mick Jagger, which is  $G$ . Initially, it may sound odd to say that we can know this sort of thing a priori; but imagine what you would say if you found out that it is actually true that one thing is  $F$ , and something else is  $G$ . It seems plausible that you would say that one has as much claim to be Mick Jagger as the other, so that if one is Mick Jagger and the other isn't, that must be because there's some other relevant difference between them. So suppose instead you found that it is actually true that there is an  $F$  which is Mick Jagger. Could there then be some other object which is  $G$ , which is *not* Mick Jagger, and which otherwise differs from the  $F$  in no relevant respect? By the above line of reasoning, it seems not. But this just goes to show that it is plausible, after all, that upon learning that the antecedent of the above conditional is actually true, I can know, without knowing anything else about the world, that its consequent must also be true. And this means that it is plausible, after all, that the relevant conditional is a priori. Since this conditional is equivalent to [10], this also shows that [10] is a priori.

If this line of reasoning were convincing, what it would show is that sentences like [10] are such that, whatever non-name-involving information we are given about a scenario, we invariably would judge [10] to be true of that scenario. But the slide from this claim to the conclusion that [10] is a priori is far from innocent and indeed is, in a way, precisely what is at issue. The question is whether we can conclude, on the basis of the claim that a sentence is epistemically necessitated by every (neutrally described) scenario, that it is a priori. If you think that claims [A]-[B] about a priori knowledge are plausible, it is this inference — and hence the epistemic two-dimensionalist way of thinking about a prioricity — that you should reject.

This is a crucial point. To determine whether some material conditional  $\lceil p \rightarrow q \rceil$  is a priori, we ask:

- (i) Does that knowledge that  $p$  is the case put one in a position, without any further information, to know that  $q$  is the case?



To determine whether such a material conditional has a necessary epistemic intension, we ask:

- (ii) Is any information, stated without use of names or other non-semantically-neutral expressions, which is sufficient to know that  $p$  is the case also sufficient to know that  $q$  is the case?

It should be uncontroversial that (i) and (ii) are distinct questions, and that the project of understanding a prioricity as necessity of epistemic intension can succeed only if (i) and (ii) always get the same answer.

The sentences considered above are, in effect, examples of material conditionals which are such that the answer to question (i) appears to be ‘No.’ To see why this seems to be the right verdict about [10], imagine that you come to know that Mick Jagger is  $F$ , but not  $G$ , and ask whether you are then in a position to know, without any further information, that nothing else is  $G$ . It seems that you are not, because it seems clear that you might entertain the possibilities that some non-Mick thing is  $G$ , and that nothing is  $G$ , and say of each: ‘for all I know, that’s the way things are.’ This is strong evidence that the answer to question (i), asked about [10], is ‘No.’

But if the argument above is correct, then the answer to question (ii), asked about [10], is ‘Yes’: [10] has a necessary epistemic intension. If this is right, then [10] is a plausible example of a sentence with respect to which questions (i) and (ii) deliver different answers, and hence a plausible counterexample to the view that a prioricity is possession of a necessary epistemic intension. One might, of course, want to find a flaw in the argument that the answer to question (ii) in this case is ‘Yes’, or resist the view that the answer to question (i) in this case is ‘No.’ The present point is just that one cannot plausibly argue that the answer to question (i) in this case is ‘Yes’ *because* the answer to question (ii) in this case is ‘Yes.’ After all, the possibility of there being a divergence between answers to these two questions just is the main point of difference between opponents and proponents of epistemic two-dimensionalist views of the a priori.

*Objection 2: Even if there is some way of blocking this argument, the epistemic two-dimensionalist’s explanation of the epistemic difference between [1] and [2] fails the non-circularity condition.*

One of the notions used in defining epistemic intensions was epistemic necessitation. So, if we are to give a noncircular explanation of a prioricity in terms of necessity of epistemic intension, we had better be able to explain epistemic necessitation in some way which does not presuppose facts about what is and what is not a priori. But we already know from our discussion above that, given the Core Thesis, the following biconditional must be necessarily true:

$e$  epistemically necessitates  $S$  iff  $(e \rightarrow S)$  is a priori.

This points up the difficulty in giving an account of epistemic necessitation which does not presuppose facts about the a priori: it is equivalent to giving an account of a priori

entailment which does not presuppose facts about the a priori. But this means that epistemic two-dimensionalism offers an explanation of a prioricity if and only if it is packaged with an independent account of a priori entailment — which, given that a prioricity can be defined in terms of a priori entailment, is an explanatory virtue which any theory has.

So (supposing, still, that we have a response to the Mick Jagger argument) epistemic two-dimensionalism can't explain what it is for a sentence to be a priori, or what a prioricity consists in. However, for all this, the epistemic two-dimensionalist might suggest that epistemic intensions provide a useful framework for modeling epistemic possibility — much as possible worlds and secondary intensions provide a useful framework for modeling metaphysical possibility, even if the framework of possible worlds, on many views, builds in facts about what is and is not possible.

Two points about this:

1. Non-two-dimensionalists might be able to provide the same sort of framework with epistemically possible but metaphysically impossible worlds, so it is not obvious that this is a distinctive advantage of two-dimensionalist semantics.
2. By itself, giving a way of classifying sentences as a priori which assumes facts about which sentences are and are not a priori is not especially impressive. It would be more impressive if the semantic values used in the classification — the epistemic intensions of sentences and sub-sentential expressions — could help solve some of the other puzzle cases for the Millian-Russellian discussed above. In that case, we would have reason to think that epistemic intensions are, genuinely, a kind of meaning, and hence that we had found some connection between a prioricity and meaning of the sort the two-dimensionalist aims to establish.

### *3.2 Cognitive significance*

As noted above, one version of Frege's puzzle which is challenging for the Millian-Russellian is to explain the difference in cognitive significance between sentences like [1] and [2]:

[1] If Hesperus exists, then Hesperus is Hesperus.

[2] If Hesperus exists, then Hesperus is Phosphorus.

Intuitively, the first sentence says something significant: it expresses something which one might well discover. The second, by contrast, seems to express a triviality; it seems to lack this sort of cognitive significance. This seems puzzling because the two sentences are alike the the level of secondary intensions. It is, therefore, extremely tempting to think that differences in cognitive significance like that exemplified by this pair of sentences might be explained by differences in epistemic intension.

This can seem very plausible when we look at examples like the one above. Below we'll have the opportunity to discuss the epistemic intensions of names in more detail; but for now let's assume that 'Hesperus is Phosphorus' is a posteriori, so that 'Hesperus is not Phosphorus' is epistemically possible. 'Hesperus is not Phosphorus' will then be epistemically necessitated by at least one scenario, so that 'Hesperus is Phosphorus' will

be false at that scenario. By contrast, ‘Hesperus is Hesperus’ is a priori — so, given the argument for the Core Thesis given above, it will be true at every scenario.<sup>3</sup>

Hence epistemic two-dimensionalism promises to explain the difference in cognitive significance between [1] and [2] in terms of the difference in epistemic intension between the two names.

However, any two expressions can be known a priori to be coreferential will have the same epistemic intension, and this poses a problem for the proposed explanation of differences in cognitive significance, because we can generate a difference in cognitive significance via the substitution of two such terms, as in the following:

[14] 49 = 49.

[15] 49 = the positive square root of 2401.

[1] and [2] seem to differ in cognitive significance in just the same way as do [14] and [15]. For this reason, epistemic intensions do not by themselves offer a solution to Frege’s puzzle.

Chalmers has suggested two kinds of responses to this problem: (1) appealing to structured epistemic intensions to explain the difference between [14] and [15]; and (2) introducing a third dimension of meaning to explain the difference. I will discuss these in turn.

### 3.2.1 *Structured epistemic intensions*

One idea is that we can solve this problem by moving from epistemic intensions as defined above to more fine-grained structured epistemic intensions, thereby finding a difference between pairs of expressions, like ‘49’ and ‘the positive square root of 2401’, which can be known a priori to be coreferential, on the grounds that the latter but not the former is semantically complex (see Chalmers (forthcoming), §3). This is a fair point about the above example, but does not in the end help very much. The problem is that the explanation of cognitive significance in terms of semantic complexity and simplicity seems to overgenerate. Consider the following sentence:

[16] 49 = the number 49.

This seems to lack cognitive significance, and to be in that respect on par with [14]. But on the proposal that we explain cognitive significance not by difference in epistemic intension but by difference in structured epistemic intension, [16] gets lumped in with [15]. This is surely a mistake. The notion of ‘cognitive significance’ is already notoriously difficult to pin down; if we are to claim that [14] but not [16] lacks cognitive significance, and that [16] like [15] has cognitive significance, it is far from clear that any intuitive distinction at all is being captured.

---

<sup>3</sup>Here and elsewhere, I’m simplifying by forgetting about the possibility that Hesperus might not exist.

### 3.2.2 Structured epistemic intensions + logical truth

A different idea is to say that the cognitively insignificant propositions are those expressed by sentences which can be turned into logical truths by substitution of terms with the same epistemic intension. This would classify [16] as insignificant, which is good; but would classify all logical truths this way, which is bad.

However, perhaps the epistemic two-dimensionalism can use some account like this to capture a notion of analyticity which might be useful. (But see the discussion of the possibility of using distinct names with the same epistemic intension below for some reasons to doubt this.)

### 3.2.3 Three-dimensionalism?

A different idea is that the flexibility of the epistemic two-dimensionalist framework might come to our aid. (This is suggested in Chalmers (ms.), §7.) Just as we defined epistemic intensions to track the epistemic properties of sentences, so we can define a *third* dimension of meaning which will more closely correspond to Fregean senses in tracking differences in cognitive significance. Epistemic intensions are defined in terms of a kind of idealized a priori knowability; but we could define *Frege-intensions* in terms of some other property more closely tied to the differences in cognitive significance that Frege's puzzle tracks. Let's introduce 'obvious' as a term of art for the property of lacking cognitive significance, in the sense sketched above. We could then use this property to define Frege-scenarios (all those possibilities which are such that it is not obvious that they are not actual) Frege-necessitation (obvious entailment), and thereby a new kind of semantic value, Frege-intensions. Perhaps we can explain differences in cognitive significance using Frege-intensions, rather than epistemic intensions.

*Objection 1: this fails non-circularity, for the same reason as the epistemic two-dimensionalist explanation of the apparent epistemic difference between [1] and [2].*

*Objection 2: it is not even clear that we have succeeded in defining a coherent notion.*

It seems clear that the property of being an 'obvious' entailment in the above sense will not be transitive; there can be a trio of sentences  $p, q, r$  such that  $\ulcorner p \rightarrow q \urcorner$  lacks cognitive significance,  $\ulcorner q \rightarrow r \urcorner$  lacks cognitive significance, but  $\ulcorner p \rightarrow r \urcorner$  has cognitive significance, and so is *not* obvious. To see this, it is sufficient to note the possibility of proofs which are such that each step in the proof is obvious despite the fact that, prior to seeing the proof, the relation between the premises and the conclusion is far from obvious.

But this, plus the definition of Frege's intensions, leads to incoherence. If  $\ulcorner p \rightarrow q \urcorner$  lacks cognitive significance, then every scenario at which the Frege-intension of  $p$  determines truth is one at which the Frege-intension of  $q$  does as well; and if  $\ulcorner q \rightarrow r \urcorner$  lacks cognitive significance, then every scenario at which the Frege-intension of  $q$  determines truth is one at which the Frege-intension of  $r$  does as well. Put these together, and what you get is that every scenario at which the Frege-intension of  $p$  is true is one at which the Frege-intension of  $r$  is true; but this contradicts our supposition that  $\ulcorner p \rightarrow r \urcorner$  could have cognitive significance, and hence be false at at least one scenario.

Suppose we grant this point, and set aside Frege-intensions. Can't the epistemic two-dimensionalist still offer an explanation of *some* examples of differences in cognitive significance in terms of epistemic intensions? It may be the case that some a priori truths are cognitively significant, but many are not; and for the latter category, perhaps we can give at least a partial explanation of their lack of cognitive significance in terms of the necessity of their epistemic intension.

In a way, this kind of claim is unobjectionable — but it is important to see here that two-dimensionalism is playing no explanatory role. All we need for the proposed partial explanation of cognitive significance (or the lack thereof) is the claim that some examples of sentences which lack cognitive significance are also examples of the a priori; and this is presumably a claim that anyone who is willing to talk about sentences being a priori and having or lacking cognitive significance will endorse. Two-dimensionalism could, of course, deepen the proposed explanation by explaining a prioricity; but we have already seen that two-dimensionalism offers no such explanation.

Chalmers says that there are various ways of dealing with non-transitivity; if so, there must be some way of blocking the foregoing argument.

### 3.3 *Substitution failures*

Let's now examine epistemic two-dimensionalism's capacity to explain the apparent differences in truth-value between sentences like [3] and [4]:

[3] John believes that Hesperus is visible in the evening.

[4] John believes that Phosphorus is visible in the evening.

The epistemic two-dimensionalist's strategy for providing a difference in truth value between these two sentences will be to provide a semantics for attitude ascriptions which makes their truth conditions sensitive to the epistemic intensions, as well as the secondary intensions, of expressions in the complement of the ascription.

#### 3.3.1 *A simple semantics for attitude ascriptions*

The simplest way of constructing such a semantic theory will make it an instance of the following schema:

‘ $a$  believes that  $S$ ’ is true in a context  $C$  iff  $a$  has a belief which (i) has the same secondary intension as  $S$  in  $C$ , and (ii) has a epistemic intension which bears  $R$  to the epistemic intension of  $S$  in  $C$ <sup>4</sup>

---

<sup>4</sup>An account like this is suggested in, among other places, Chalmers (2002b) and Chalmers (2006). One point worth noting about this proposal is that it implies that mental states, like sentences, are the kinds of things that can have primary and secondary intensions. This seems to imply that mental states are underwritten by bearers of content which are at least to that extent analogous to sentences. While this is a substantial commitment, I think that most two-dimensionalists would be happy with it.

instances of which are obtained by replacing ‘ $R$ ’ with a term for some similarity relation, which might depend on context, between epistemic intensions.

A pleasingly simple instance of this schema would let ‘ $R$ ’ stand for the identity relation. However, indexicals make this suggestion unworkable; one way to see this is to note that if I express a belief with the words ‘I am home’, it is impossible for you to report that belief using a sentence which has both the same secondary and epistemic intensions as the belief. But it shouldn’t, in all contexts, be impossible to report such a belief truly.<sup>5</sup> Instead we should take ‘ $R$ ’ to be some sort of similarity relation, which may depend on the standards which are (in some sense) operative in the context.

A first point to make about this idea is familiar from the preceding section: any such explanation of substitution failures will be incomplete, because we can generate phenomena similar to [3] and [4] using substitution of singular terms which can be known a priori to be coreferential, and so must have the same epistemic intension. For example, consider the following pair of sentences:

[17] John believes that 7 is odd.

[18] John believes that the positive square root of 49 is odd.

Just as [3] and [4] appear to differ in truth-value, so we can imagine cases in which [17] and [18] differ in truth value. Insofar as these appear to be instances of the same phenomenon, the two-dimensionalist’s solution here looks worryingly partial.

### 3.3.2 *The semantics once-complicated*

But as above, we can perhaps get around this by appeal to structured epistemic intensions, giving a semantics of the following sort:

‘ $a$  believes that  $S$ ’ is true in a context  $C$  iff  $a$  has a belief which (i) has the same secondary intension as  $S$  in  $C$ , and (ii) has a structured epistemic intension which bears  $R$  to the structured epistemic intension of  $S$  in  $C$

A problem for any account of this sort arises from the fact that the epistemic intensions of names will vary widely between speakers.

Imagine (what is in fact true) that I can never remember which name is supposed to be linked to the description ‘the morning star’ and which is supposed to be linked to ‘the evening star.’ In this case, it is plausible to think that the epistemic intensions of the names converge for me: given any scenario, I will identify the denotation of a description ‘the  $F$ ’ as ‘Hesperus’ if and only if I identify it as ‘Phosphorus’. Since the two names have the same primary and secondary intensions out of my mouth, then, given that some version of the semantics for attitude ascriptions sketched above is true, then, out of my mouth, the two names *are* substitutable *salva veritate* in attitude ascriptions.

---

<sup>5</sup>For arguments against the view that attitude ascriptions require identity of epistemic intension, along with related views, see Soames (2005).

[Could the epistemic two-dimensionalist deny that in this situation the names really do have the same epistemic intension for me, on the grounds that *ceteris paribus* I would identify the object as the referent of ‘Hesperus’ which speakers in the scenario refer to as ‘Hesperus’? Not unless he is willing to say that I know a priori facts about the names of objects.]

This leads to two related sets of problems.

1. The first is simple. Since the two names have the same epistemic intensions out of my mouth, all of the following must be true, as uttered by me:

Hammurabi believed that Hesperus is Phosphorus.

Anyone who has ever believed that Hesperus is visible in the evening has also believed that Phosphorus is visible in the evening.

It is knowable a priori that if Hesperus exists, Hesperus is Phosphorus.

But these are the kind of counterintuitive assignments of truth-values which epistemic intensions were meant to block.

2. The two-dimensionalist might reply that epistemic intensions don’t explain why these sentences seem false when uttered by me, but they do at least explain why sentences of this sort seem false when uttered by you, for whom the epistemic intensions of the names diverge. But the fact that the epistemic two-dimensionalist treats our utterances differently in this way is not a strength of the view.

Any plausible semantics for attitude ascriptions should not, given the above description of the case, rule out my being able to report the thoughts you express using the word ‘Hesperus’ using the same word, and the same for ‘Phosphorus.’ So let’s suppose that the epistemic intensions we assign to the word are ‘similar enough’ to allow for this.

Now let’s imagine that a friend of ours uses the two terms with the same epistemic intensions as you, and, furthermore, that this friend has quite radical astronomic views. Suppose he turns to us and says, ‘Hesperus is not Phosphorus!’ Since we know that our friend assigns these names the same epistemic intension you do, and that our epistemic intensions are similar enough for me to report your speech, I can truly say of our friend,

[19] He said that Hesperus is not Phosphorus.

and hence, since the names have the same epistemic intension for me, if [19] is true out of my mouth, I can also truly say

[20] He said that Hesperus is not Hesperus.

But given that you can report my speech, you can report my utterance of [20] by saying

[21] JS said that he said that Hesperus is not Hesperus.

Moreover, since you might know that my utterance was true, you might expand upon [21] by saying

[22] JS said that he said that Hesperus is not Hesperus, and what JS said was true.

and it seems that your utterance of [22] will be true. But if you can truly utter [22], then you must also be able to truly utter [20]. But [20] certainly seems to be

false. So it seems that the epistemic two-dimensionalist has to assign [19] and [20] the same truth-value out of both of our mouths.

Another way to see the oddness here is by imagining our friend with radical astronomical views uttering

[23] I said that Hesperus is not Hesperus.

This certainly seems as though it should be false; the problem with Millianism is supposed to be that it is forced to say that utterances such as this are, even though the speaker takes them to be false, really true. But the epistemic two-dimensionalist is forced into saying that [23] is true as well — after all, she takes your utterance of [20] to be true, and you and our friend with radical astronomical views assign ‘Hesperus’ the same epistemic intension.

Can the epistemic two-dimensionalist plausibly reply to this case by saying that the contextually defined relation  $R$  shifts during the conversation in such a way that some steps in the argument are blocked?

### 3.3.3 *The semantics twice-complicated*

It seems that the epistemic two-dimensionalist must respond to this sort of argument by saying that the truth-values of attitude ascriptions are sensitive, not just to the epistemic and secondary intensions of the expressions in the complement of the ascription, but also to the words used in the ascription. That is, she might hold that the right semantics for attitude ascriptions is some version of the following:

‘ $a$  believes that  $S$ ’ is true in a context  $C$  iff  $a$  has a belief which (i) has the same secondary intension as  $S$  in  $C$ , (ii) has a structured epistemic intension which bears  $R$  to the structured epistemic intension of  $S$  in  $C$ , (iii)  $S$  bears some relation  $R^*$  to the sentence that the subject of the belief uses to express the belief.

One can of course say this; but one needs to say something substantive about what the two key relations  $R$  and  $R^*$  are. One can’t just say that they are defined by our intuitions about the truth conditions of attitude ascriptions without falling afoul of the noncircularity constraint. There is a difference between using speaker intuitions as data for which a semantic theory must account — which is of course legitimate — and using speaker intuitions to define semantic contents which are then presented as a validation and explanation of those very intuitions.

### 3.3.4 *A fallback position*

Could we abandon the semantics suggested above in favor of one given solely in terms of secondary intensions, and use the epistemic intensions solely to explain the ‘pragmatics’?<sup>6</sup> The suggestion is initially a plausible fall back position, since then, it might seem, the semantics of the epistemic two-dimensionalist is no worse off than his opponent, but builds

<sup>6</sup>This is suggested as a possibility in Chalmers (2002a), p. 631 note 23.



in a kind of mechanism for generating the relevant pragmatic facts about uses of these sentences.

But, on closer examination, this suggestion is not very promising; what is pragmatically conveyed by attitude ascriptions typically has a lot to do with the intentions of parties to the relevant conversation, and very little to do with the epistemic intensions of the relevant expressions for the speaker. Consider again the example above, where I do and you do not associate the same epistemic intension with ‘Hesperus’ and ‘Phosphorus.’ Despite this, we pragmatically convey the same thing by our utterances of ‘He said that Hesperus is not Phosphorus.’ And the same goes for other such cases. If someone says ‘Lois does not know that Clark Kent is Superman’, we know what they are trying to get across without investigation into the epistemic intensions that they associate with the two names.

### 3.4 *The problem of the essential indexical*

The problem of the essential indexical is often thought of as the problem of explaining the apparent difference between sentences like [5] and [6]. So put, there are really at least three different versions of the problem, depending on which apparent difference between these sentences we focus on:

1. I can believe that Jeff Speaks is  $F$  without believing that I am  $F$ , even though I am Jeff Speaks.
2. I can believe that Jeff Speaks is  $F$  without being in a position to deduce a priori that I am  $F$ , even though I am Jeff Speaks.
3. My beliefs that I have various properties have a kind of direct effect on my actions which my beliefs that Jeff Speaks has various properties do not.

Version (1) of the problem is just an instance of the more general problem of apparent substitution failures involving coreferential terms in attitude ascriptions; we’ve already seen that the two-dimensionalist does not have much helpful to say about this.

Version (2) of the problem is another way of saying that certain conditionals of the form

$$n \text{ is } F \rightarrow \text{I am } F$$

are not a priori for a speaker, even when the speaker is the referent of ‘ $n$ .’ As such, it is a fact about which sentences are, and which are not, a priori; and we’ve already seen that epistemic two-dimensionalism builds in rather than explains such facts.

Versions (1) and (2) of the problem are just special cases of the problems of explaining the a priori and apparent substitution failures in attitude ascriptions; so it remains to discuss only version (3). The third version, or aspect, of the problem of the essential indexical is the most interesting. What is it about the belief that I would express by ‘I am on fire’ which gives it a kind of direct connection to action which the belief that I would express by ‘Jeff Speaks is on fire’ might lack?

Consider first how the *contextual* two-dimensionalist would answer this question. The contextual intension of ‘I am on fire’ is something like the set of contexts in which the

speaker of the sentence is on fire. So, to utter this sentence is to say something with the contextual intension that the very person saying ‘I am on fire’ is on fire. This seems to capture a sense in which the belief expressed by ‘I am on fire’ is essentially self-directed in a way in which a belief expressed by ‘*n* is on fire’ needn’t be; so it might seem as though the contextual two-dimensionalist has the makings of an explanation of (3) above.

But matters are different with epistemic two-dimensionalism. Whether we think of scenarios according to the metaphysical or the epistemic conception, they Recall that scenarios are complete hypotheses about the world which cannot be ruled out a priori. Because these hypotheses are supposed to describe ‘centered’ hypotheses, they will contain indexicals like ‘I.’<sup>7</sup> Let’s suppose that it is epistemically possible (for me) that the following are both true:

Scenario 1  
 Jeff Speaks is on fire.  
 I am not on fire.

and that it is epistemically possible that the following are both true:

Scenario 2  
 Jeff Speaks is not on fire.  
 I am on fire.

Then the epistemic intensions of ‘Jeff Speaks is on fire’ and ‘I am on fire’ will differ — which is, from the epistemic two-dimensionalist’s point of view, all to the good.

But of course, to make room for these two sentences to have different epistemic intensions is not to offer any solution to version (3) of the problem of the essential indexical: that problem asks us not to explain how these two sentences could differ in content, but why ‘I am on fire’ has a kind of built-in direct connection to action which ‘Jeff Speaks is on fire’ does not. And it seems fairly clear that epistemic two-dimensionalism has nothing to say about this. Within the framework of epistemic two-dimensionalism, ‘I’ is just one term among others which can appear in scenarios; so far as epistemic intensions go, ‘I’ is indistinguishable from any singular term whose reference varies from context to context. Nothing in the framework explains why ‘I am *F*’ should have any closer connection to action than ‘*n* is *F*.’

The two-dimensionalist might reply as follows:

If I learned that Scenario 2 is actual, then I would be disposed to douse myself with water; if I learned that Scenario 1 is actual, then I wouldn’t be. But (to oversimplify a bit) epistemic contents just are sets of scenarios; so the difference between the epistemic intensions of ‘I am on fire’ and ‘Jeff Speaks is on fire’ is the difference between Scenarios 2 and 1. But then the fact that

---

<sup>7</sup>This is not explicit in Chalmers (2006). However, the discussion in §3.4.1 makes it clear that corresponding to one ‘objective’ — i.e., indexical-free — description of the world, there will typically be many scenarios. My thought here is that to get this result — which the two-dimensionalist needs — indexical expressions need to be a part of scenario descriptions.

Scenario 2 but not Scenario 1 would dispose me to douse myself with water shows that the epistemic intension of ‘I am on fire’ but not ‘Jeff Speaks is on fire’ has an immediate connection to action. Doesn’t this show that epistemic intensions do solve the third version of the problem of the essential indexical?

I don’t think so. Suppose we ask: ‘Why would learning that Scenario 2 is actual lead me to douse myself with water?’ Presumably, because this scenario includes the claim, ‘I am on fire.’ Given this, we can hardly turn around and answer the question ‘Why does “I am on fire” have a direct connection to action?’ by saying ‘Because it has an epistemic intension which rules out Scenario 1 in favor of Scenario 2.’

Contextual two-dimensionalism promised an answer to the problem of the essential indexical because contextual intensions are, in a sense, inherently indexical. But epistemic intensions are not inherently indexical – except in the uninteresting sense that indexical expressions are part of the scenarios which epistemic intensions take as their arguments.<sup>8</sup>

## 4 Conclusion

To conclude: epistemic two-dimensionalism can do no work in explaining which sentences are a priori and which are not, and can’t solve either version of Frege’s puzzle or the problem of the essential indexical.

The foregoing argument not show that epistemic two-dimensionalism is incoherent, or that expressions do not have epistemic intensions. But it at least raises the question whether epistemic intensions are properties of expressions which, once discovered, can just as well be forgotten.

But isn’t there a further motivation for epistemic two-dimensionalism, beyond the capacity of epistemic intensions to play various explanatory roles — namely, the aim of restoring, in Chalmers’ words, ‘the golden triangle of constitutive connections between meaning, reason, and modality’? Given the foregoing, it is, at best, extremely obscure how epistemic intensions could play even this role. To show that there is a substantive (let alone constitutive) connection between reason and modality, we need to do more than show that we can find a class of objects (whether sentences or the states underlying propositional attitudes) and define one class of properties of those objects in terms of reason, and another class of properties of those objects in terms of modality. The unity of the bearer of the two classes of properties hardly provides a substantive connection between the properties themselves.

It is also, I suggest, somewhat obscure why we should want the ‘golden triangle’ to be restored. It is a traditional idea, for example, that some predications are necessary because of the nature of the subject of the predication. Why should it be surprising that the natures of things fail, in the standard case, to stand in any constitutive relation to human thought?

---

<sup>8</sup>While the idea that epistemic intensions can solve the problem of the essential indexical figures prominently in Chalmers (2002a), it gets no mention in later papers like Chalmers (2006). Perhaps this indicates that Chalmers would now agree with the foregoing.

## References

- David Chalmers, 2002a. The Components of Content. In *Philosophy of Mind: Classical and Contemporary Readings*, edited by David Chalmers, 608–633. New York: Oxford University Press.
- David Chalmers, 2002b. On Sense and Intension. *Philosophical Perspectives* 16.
- David Chalmers, 2004. Epistemic Two-Dimensional Semantics. *Philosophical Studies* 118(1-2):153–226.
- David Chalmers, 2006. The Foundations of Two-Dimensional Semantics. In *Two-Dimensional Semantics*, edited by Manuel García-Carpintero and Josep Macià, 55–140. New York: Oxford University Press.
- David Chalmers, forthcoming. Propositions and Attitude Ascriptions: A Fregean Account. *Noûs* .
- David Chalmers, ms. The Nature of Epistemic Space. [consc.net/papers/espace/html](http://consc.net/papers/espace/html) .
- Nathan Salmon, 1989. How to Become a Millian Heir. *Noûs* 23:211–220.
- Scott Soames, 2005. *Reference and Description: The Case Against Two-Dimensionalism*. Princeton, NJ: Princeton University Press.
- Scott Soames, 2007. Actually. *Aristotelian Society Supplement* LXXXI:251–277.

*Proof of equivalence of epistemic necessitation and a priori consequence.*

Given the Core Thesis, epistemic necessitation must be some relation  $R$  such that for any sentence  $S$ ,  $S$  is a priori iff for all scenarios  $e$ ,  $R(e, S)$ . What we want to show is that, given this,

$$R(e, S) \text{ iff } \ulcorner e \rightarrow S \urcorner \text{ is a priori.}$$

One way to show this is to begin with the following Lemma:

$$R(e, S) \text{ iff } R(e, e \rightarrow S)$$

This Lemma may be intuitively obvious, but we can also give the following argument for it, assuming only that the Core Thesis is true.

Left to right:  $\ulcorner S \rightarrow (e \rightarrow S) \urcorner$  is a priori, so, given the Core Thesis, it must be true at every scenario. So any scenario which epistemically necessitates  $S$  must also epistemically necessitate  $\ulcorner e \rightarrow S \urcorner$ .

Right to left: if  $e$  epistemically necessitates  $S$ , then knowing that  $e$  is the case puts one in a position to know that  $\ulcorner e \rightarrow S \urcorner$  is true. There are only two ways for  $\ulcorner e \rightarrow S \urcorner$  to be true: for  $e$  to be false, or for  $S$  to be true. But knowing that  $e$  is the case can never put one in a position to know that  $e$  is false; so if  $e$  epistemically necessitates  $\ulcorner e \rightarrow S \urcorner$ ,  $e$  must put one in a position to know that  $S$ . But then  $e$  epistemically necessitates  $S$ .

We then argue for the equivalence of a priori entailment and epistemic necessitation, i.e. the truth of the biconditional

$$R(e, S) \text{ iff } \ulcorner e \rightarrow S \urcorner \text{ is a priori.}$$

as follows:

Left to right: suppose that  $\ulcorner e \rightarrow S \urcorner$  is not a priori. Then, by the Core Thesis, there must be some scenario  $e^*$  such that  $R(e^*, e \rightarrow S)$ . Either  $e^* = e$  or  $e^* \neq e$ . But the latter is impossible, since, scenarios being complete epistemic possibilities, any scenario must epistemically necessitate the negation of any other scenario; and if  $e^*$  did epistemically necessitate  $e$ , then, given that  $\ulcorner e \rightarrow (e \rightarrow S) \urcorner$  is a priori,  $e^*$  would epistemically necessitate  $\ulcorner e \rightarrow S \urcorner$ , which is contrary to our supposition. So  $e^* = e$ , and  $R(e, e \rightarrow S)$ . But then it follows from our Lemma that  $\neg R(e, S)$ .

Right to left: suppose that  $\ulcorner e \rightarrow S \urcorner$  is a priori. Then, by the Core Thesis, for any scenario  $e$ ,  $R(e^*, e \rightarrow S)$ ; so, in particular,  $R(e, e \rightarrow S)$ . But then, by our Lemma,  $R(e, S)$ .