Ambitious Two-Dimensionalism

by

Scott Soames

School of Philosophy

USC

To Appear in

On Sense and Direct Reference:

A Reader in Philosophy of Language

Matthew Davidson, editor

McGraw-Hill

Ambitious Two-Dimensionalism
Scott Soames

Three decades ago, a group of philosophers led by Saul Kripke, Hilary Putnam, and David Kaplan ushered in a new era by attacking presuppositions about meaning that occupied center stage in the philosophy of the time. Among these presuppositions were the following:

(i)   The meaning of a term is never identical with its referent. Instead, its meaning is a descriptive sense that encodes conditions necessary and sufficient for determining its reference.

(ii)   Understanding a term amounts to associating it with the correct descriptive sense. Different speakers who understand a predicate of the common language, or a widely used proper name such as *London*, associate it with essentially the same sense. However, for many names of lesser-known things, the defining descriptive sense associated with the name varies from one speaker to the next.

(iii)   Since the meaning of a word, as used by a speaker s, is the descriptive sense that s mentally associates with it, meaning is transparent. If two words mean the same thing, then anyone who understands both should be able to figure that out by consulting the sense that he or she associates with them. For similar reasons, word meanings and mental contents are entirely dependent on factors internal to speakers.

 (iv)   Apriori and necessary truth amount to the same thing. If they exist at all, both are grounded in meaning.

 (v)   Claims about objects having or lacking properties essentially -- independent of how they are described -- make no sense. Even if a term t designates o and *Necessarily t is F* is true, there will always be another term t* designating o for which *Necessarily t* is F* is false. Since it would be arbitrary to give either sentence priority in determining the essential properties of o, the idea that objects have, or lack, properties essentially, must be relativized to how they are described.

(vi)   Since the job of philosophy is not to come up with new empirical truths, its central task is that of conceptual clarification, which proceeds by the analysis of meaning.

These doctrines and their corollaries provided the framework for much of the philosophy in the analytic tradition prior to the 1970's. Of course, not every analytic philosopher accepted all tenets of the framework, and some, like Quine, rejected the traditional notions of meaning, necessity,

and apriority altogether. However, even Quine -- the framework's most severe critic – believed that if the traditional notions make sense at all, then they must be related more or less along the lines indicated. What was, for the most part, absent was a recognition that all of these notions do make sense, and are important for philosophy, even though they are mischaracterized by the traditional framework.

That changed with Kripke, Putnam, Kaplan, and the line of research growing out of their work. Today, all doctrines of the framework have been challenged, and replacements have been suggested. However, no new consensus has been reached. Although everyone recognizes the need to take into account the arguments of Kripke and his fellow anti-descriptivists, some continue to believe that the traditional paradigm contained much that was correct, and that a new, more sophisticated version of descriptivism should be put in its place. Even those who reject the idea of a descriptivist revival, and want to push the anti-descriptivist revolution further, have found the task of constructing a positive, non-descriptivist conception of meaning to be daunting. In short, the struggle over the legacy of the original challengers to descriptivism is far from over.

Here, I will discuss an important part of that struggle. In the last 25 years a systematic strategy has grown up around a technical development called *two-dimensional modal logic*, for reviving descriptivism, reconnecting meaning, apriority, and necessity, and vindicating philosophy as conceptual analysis along recognizably traditional lines. Although the logical and semantic techniques are new, the motivating ideas are old. Since many of these ideas were not

without plausibility, it is not surprising that an attempt has been made to reinstate them.   But there is more to the attempted revival than this. Anti-descriptivism has brought with it problems of its own. I will begin by explaining how these problems have motivated a cluster of views I call *ambitious two-dimensionalism*. I will then identify what I take to be the shortcomings of these views, and explain why I believe that no version of ambitious two-dimensionalism can succeed.

The Anti-Descriptivist Revolution

First a word about the anti-descriptivist revolution. In <u>Naming and Necessity</u>, Kripke offers two main arguments against the view that the meanings of names are given by descriptions associated with them by speakers, plus a further class of arguments against the view that the reference of most names is fixed in this way.[1] The modal argument holds that since sentences containing proper names are true in different possible world-states than corresponding sentences containing descriptions, names can't mean the same as descriptions. The epistemological argument holds that if names really meant the same as descriptions, then certain sentences containing them would express apriori truths. Since these sentences don't express such truths, names aren't synonymous with descriptions. Finally, Kripke's semantic arguments show that the descriptions speakers would volunteer in answer to the question ***To whom, or what, do you use 'n' to refer?*** sometimes fail to denote the object to which n really refers, and sometimes denote something to which n doesn't refer.[2] From this he concludes that, except in relatively rare cases, the linguistic rule mastered by speakers for determining the referent of a name is not that it is to refer to whatever is denoted by a set of associated descriptions. Instead, reference is determined by

---

[1] Saul Kripke, <u>Naming and Necessity</u>, (Cambridge MA: Harvard University Press), 1980; originally given as lectures in 1970, and originally published in 1972.

[2] Here, and throughout, boldface italics are used as corner quotes.

historical chains of reference transmission connecting later uses to earlier ones, and ultimately to initial baptisms introducing the name. A similar story is told for natural kind terms.

To this David Kaplan added an account of indexicals.[3] On this account, to know the meaning the first person, singular pronoun, is to know that one who uses it in a sentence -- ***I am F*** – refers to oneself, and says of oneself that one "is F". Similar rules govern other indexicals. These rules both tell us how the referents of indexicals depend on aspects of contexts in which they are used, and implicitly identify the semantic contents of indexicals with their referents. What is semantic content? The semantic content of a sentence is the proposition it expresses. Sentences containing indexicals express different propositions, and so have different contents in different contexts. Nevertheless, the meaning of such a sentence is constant; it is a function from contexts to contents. Kaplan's word for this is *character*. The picture is recapitulated for subsentential expressions. The character of the pronoun *I* is a function that maps an arbitrary context C onto the agent of C, which is its semantic content of *I* in C.

There are two anti-descriptivist implications here. First, the referents of at least some indexicals are not determined by descriptions speakers associate with them. One example involves Kaplan's identical twins, Castor and Pollux, raised in qualitatively identical environments to be molecule for molecule identical and so, presumably, to associate the same purely qualitative descriptions with the same terms.[4] Despite this, each twin refers to himself, and not the other, when he uses *I*. Although this leaves open the possibility that some indexicals may have their referents semantically fixed by descriptions containing other indexicals, it precludes the possibility that all indexical reference is determined in this way.

---

[3] David Kaplan, "Demonstratives," in J. Almog, J. Perry, and H. Wettstein, eds., <u>Themes from Kaplan</u>, (New York and Oxford: Oxford University Press), 1989; originally given as lectures in 1971.

[4] "Demonstratives," p. 531.

The second anti-descriptivist implication is that since the semantic content of an indexical in a context is its referent, its content is not that of any description. The underlying picture is one in which the proposition expressed by S is a structured complex, the constituents of which are the semantic contents of the words and phrases of S. For example, the proposition expressed by the sentence *I am F* is a complex in which the property expressed by F is predicated of the agent of the context. This is the same proposition that is expressed by the formula *x is F*, relative to an assignment of o to 'x'. A similar story is told for other indexicals. As Kaplan tells it, this story has

consequences for propositional attitude ascriptions. Suppose, to adapt Russell's famous example, that on some occasion in which George IV spied Walter Scott, he gave voice to his conviction, saying *He* [gesturing at Scott] *isn't the author of* <u>Waverley</u>. Had he done this, the attitude ascription

> The author of <u>Waverley</u>, namely Scott, is such that George IV said that he wasn't the author of <u>Waverley</u>.

would have been true – as would the ascriptions

> George IV said that you weren't the author of <u>Waverley.</u> (said addressing Scott)
> George IV said that I wasn't the author of <u>Waverley</u>. (said by Scott)

On Kaplan's picture, these reports are true because the semantic content of the sentence George IV uttered (in his context), and so a proposition he asserted, is the same as the content of the complement clauses in the reports of what he said. Whatever descriptions speakers who utter these indexical sentences may happen associate with the indexicals are irrelevant to the semantic contents of the sentences they utter.   When used in attitude ascriptions, indexicals, like variables in cases of quantifying-in, are used to report an agent's attitude toward someone, or some thing, abstracting away from the manner in which the agent thinks of that person, or that thing.   All they contribute to the proposition George IV is reported as asserting is the individual Scott. We

may express this by saying that for Kaplan indexicals are not only *rigid designators*, but also *directly referential*. Some, including Nathan Salmon and me, extend this to proper names.[5]

So far, I have talked only about semantics. However, the new view of semantics is closely linked to a view that recognizes the contingent apriori and the necessary aposteriori. Since the latter will be most important for us, we will focus on it. Kripke's route to the necessary aposteriori was simple. He first used the concept of rigid designation to rebut Quine's objection to essentialism.[6] Then, with both a nondescriptive semantics and a rehabilitated conception of essentialism in place, he showed how to generate instances of the necessary aposteriori. If n is a name or indexical that rigidly designates o, and P expresses an essential property of o which is such that knowledge that o has it requires empirical evidence, then the proposition expressed by ***If n exists, then n is P*** is both necessary and knowable only aposteriori.

Reasons for Descriptivist Revival

So much for anti-descriptivism. We now turn to what some take to be the grounds for a descriptivist revival. First is the conviction that anti-descriptivists have not adequately addressed Frege's puzzle about substitution in attitude ascriptions, and Russell's problem of negative existentials. There is still a widespread belief that these problems show that names can't be directly referential.[7] Although Kripke never asserted that they were, it is hard to see how, if his doctrines are correct, they could be otherwise. According to him, the meaning of a name is never that of any description, and the vast majority of names don't even have their referents

---

[5] Nathan Salmon, <u>Frege's Puzzle</u>, (Cambridge MA: MIT Press), 1986; Scott Soames, <u>Beyond Rigidity</u>, (New York: Oxford University Press), 2002.

[6] For discussion of Quine's objection and Kripke's rebuttal see pp. 347-54 of my <u>Philosophical Analysis in the Twentieth Century, Volume 2: The Age of Meaning</u>, (Princeton: Princeton University Press), 2003.

[7] See chapter 1 of my <u>Reference and Description</u>, (Princeton and Oxford: Princeton University Press), 2005.

semantically fixed by descriptions. If these names are so thoroughly nondescriptional, it is hard to see how their meanings could be other than their referents. Consequently, one who takes this Millian view to have been refuted by Frege and Russell may naturally suspect the power of Kripke's arguments to have been exaggerated, and may be motivated to find a way of modifying descriptivism so as to withstand them.

The second factor motivating descriptivists is their conviction that critics like Kripke have focused on the wrong descriptions. To be sure, it will be admitted, for many speakers s and names or natural kind terms n, the descriptions most likely to be volunteered by s in answer to the question *To whom, or what, do you refer when using n?* neither give the meaning of n, nor semantically fix its reference. However, the referents of these terms must be determined in some way, and surely, whatever way that turns out to be is one which speakers have some awareness of, and which can be described. So, for each n, there must be some description that correctly picks out its referent -- perhaps one encapsulating Kripke's own historical chain picture of reference transmission. Some descriptivists even go so far as to suggest that descriptive theories of reference are, for all intents and purposes, irrefutable.[8] The idea is that any refutation would require an uncontroversial scenario in which n refers to some object o not satisfying the description D putatively associated with n by speakers (or in which n fails to refer to the thing that is denoted by D). However, the very judgment that n refers to o in this scenario (or doesn't refer to what D denotes) is taken by these descriptivists to demonstrate the existence of a different, implicit, description in our minds that does determine reference in the scenarios – even though we can't articulate it.

---

[8] See, for example, Frank Jackson, "Reference and Description Revisited," <u>Philosophical Perspectives 12:</u> <u>Language Mind, and Ontology</u>, 1998, especially at p.212.

The third factor motivating descriptivist revival involves the inability of some to see how any single proposition could be either both necessary and aposteriori, or both contingent and apriori, as anti-descriptivists maintain. Again, we focus on the necessary aposteriori. How, some philosophers ask, can empirical evidence about the actual world-state be required to know p, if p is true in every possible state? Surely if such evidence is required, it must have the function of ruling out possible ways in which p could be false. But if p is true in every possible world-state, then there are no such ways to rule out. So, if p is knowable at all, p must be knowable apriori. The idea that p is both necessary and knowable only aposteriori is incoherent, and any nondescriptive semantics that says otherwise must be incorrect.[9]

What gives this reasoning force is a commitment to metaphysical possibility as the only kind of possibility. On this view, there are different metaphysically possible ways the world could be, but there are no further, epistemically possible ways that the world might be. There are no world-states which, though metaphysically impossible, cannot be known by us apriori not to obtain. This restriction of epistemic possibility to metaphysical possibility renders the necessary aposteriori problematic -- since it precludes seeing it as involving metaphysically necessary propositions for which empirical evidence is needed to rule out metaphysically impossible, but epistemically possible, world-states in which they are false. When one adds to this the popular analysis of knowing p as having evidence that rules out all relevant possible ways of p's being untrue, one has, in effect, defined propositions that are both necessary and knowable only aposteriori out of existence. A different philosophical commitment that leads to the same result identifies propositions with sets of metaphysically possible world-states.[10] On this view, there is only one

---

[9] An analysis of knowledge with this consequence is given by David Lewis in "Elusive Knowledge," originally published in 1996, reprinted in his Papers in Metaphysics and Epistemology, (Cambridge: Cambridge University Press), 1999.

[10] See Robert Stalnaker, Inquiry, (Cambridge MA: MIT Press), 1984.

necessary proposition, which is known apriori. But then, the anti-descriptivist semantics that leads to the view that there are necessary aposteriori propositions must be mistaken.

The Two-Dimensionalist Strategy for Reviving Descriptivism

I now turn to the strategy of descriptivist revival, which consists of three main elements: (i) the attempt to find reference-fixing descriptions withstanding Kripke's semantic arguments, (ii) the rigidification of those descriptions to avoid the modal argument , and (iii) the use of two-dimensional semantics to explain away putative examples of the necessary aposteriori and the contingent apriori. The most popular strategy for finding reference-fixing descriptions is *causal descriptivism*, which involves extracting a description from Kripke's historical account of reference transmission. This idea is illustrated by the following passage from David Lewis.

> Did not Kripke and his allies refute the description theory of reference, at least
>
> for names of people and places?....I disagree. What was well and truly refuted
>
> was a version of descriptivism in which the descriptive senses were supposed to
>
> be a matter of famous deeds and other distinctive peculiarities. A better version
>
> survives the attack: *causal descriptivism*. The descriptive sense associated with
>
> a name might for instance be *the place I have heard of under the name*
>
> *"Taromeo"*, or maybe *the causal source of this token: Taromeo,* and for an
>
> account of the relation being invoked here, just consult the writings of causal
>
> theorists of reference.[11]

The second part of the descriptivists' strategy is to rigidify reference-fixing descriptions. The idea is to explain apparent instances of substitution failure involving coreferential names in attitude ascriptions by appealing to descriptive semantic contents of

---

[11] David Lewis, "Naming the Colors," originally published in 1997, reprinted in <u>Papers in Metaphysics and Epistemology</u>, fn. 22.

names; while using rigidification to guarantee substitution success when one such name is substituted for another in modal constructions. The final weapon in the descriptivists' arsenal is ambitious two-dimensionalism, which may be illustrated using a putative example of the necessary aposteriori.

1.   If the x: actually Fx exists, then the x: actually Fx = the x: actually Gx.

Here, we let *the F* and *the G* be contingently codesignative, non-rigid descriptions. The semantics of *actually* guarantees that since (1) is true, it is a necessary truth. Nevertheless, the knowledge reported by (2) is seen as being, at bottom, nothing over and above the knowledge reported by (3).

2.   y knows that if the x: actually Fx exists, then the x: actually Fx = the x: actually Gx

3.   y knows that if the x: Fx exists, then the x: Fx = the x: Gx

Since the latter is aposteriori, so is the former.   How can this be? [12]

        The two-dimensionalist answer is based on the relationship between the complement sentence, (1), of the ascription (2), and the complement sentence, (4), of the ascription (3).

4.   If the x: Fx exists, then the x: Fx = the x: Gx

(1) and (4) are nonequivalent in that they express different propositions, but equivalent in that they express truths in the same contexts of utterance. Since (4) expresses the same proposition p in all contexts, the two sentences express truths in all and only those contexts in which p is true.

---

[12] Although it is evident that the proposition expressed by (1) is knowable aposteriori, it is far less clear that it is knowable only aposteriori. For an argument that it is in fact possible to know it apriori (without knowing the proposition expressed by (4)), see my "Understanding Assertion," in J. J. Thomson and A. Byrne, eds., Content and Modality: Themes from the Philosophy of Robert Stalnaker, (Oxford: Oxford University Press), forthcoming. Although I will return to this point below when assessing ambitious two-dimensionalism, in presenting the view I will temporarily take it for granted that examples like (1), containing the actuality operator, are genuine instances of the necessary aposteriori.

By contrast, (1) is semantically associated with two propositions – one which it expresses in our present context, and one which states the conditions a context must satisfy if (1) is to express a truth. The former, called *the secondary proposition*, is necessary, while the latter, called *the primary proposition*, is the contingent, aposteriori truth p that is also expressed by (4).

How does this provide a two-dimensionalist explanation of the putative aposteriori status of (1)? Two main possibilities suggest themselves (between which informal discussions of two-dimensionalism often do not distinguish). The first arises from a semantic theory I call *strong two-dimensionalism*. It holds that although the proposition expressed by (1) is a necessary truth, the knowledge reported by (2) is knowledge, not of this truth, but of the conditions under which (1) expresses a truth.   On this view, there is no puzzle explaining how the proposition expressed by (1) can be both necessary and knowable only aposteriori, because it isn't. Instead, the secondary proposition associated with (1) is relevant to its modal status, while its primary proposition is relevant to its epistemic status.   All sentences that express different propositions in different contexts are seen as semantically associated with two propositions relative to any single context C: the proposition the sentence expresses in C (its secondary proposition relative to C) and the proposition that states the conditions that must be satisfied by any context in which the sentence expresses a truth (its primary proposition). The primary proposition associated with S provides the argument to the operators *it is knowable apriori* (or *aposteriori*) *that,* as well as *Jones knows (apriori* or *aposteriori) that*. The secondary proposition provides the argument to modal operators like *it is a necessary truth that*.   This is the basis for the strong two-dimensionalist's claim that every instance of the necessary aposteriori can be explained along the same lines as (1).

A slightly different explanation arises from a semantic theory I call *weak two-dimensionalism*. As before, sentences that express different propositions in different contexts are semantically associated with primary and secondary propositions. However, according to weak

two-dimensionalism, the argument provided by such a sentence to the operators *it is knowable apriori* (or *aposteriori*) *that,* and *Jones knows (apriori* or *aposteriori) that* is not its primary proposition, but its secondary proposition. On this view, (2) reports knowledge of (1)'s necessary, secondary proposition. However, this knowledge is counted as aposteriori because, it is claimed, (1)'s secondary proposition can be known only by virtue of knowing (1)'s primary proposition. As before, what makes the knowledge reported by (2) aposteriori is that (1)'s primary proposition is aposteriori. But whereas the strong two-dimensionalist claims that knowledge of the primary proposition is reported instead of knowledge of the secondary proposition, the weak two-dimensionalist claims that knowledge of the primary proposition counts as knowledge of the secondary proposition. A similar explanation is envisioned for all cases of the necessary aposteriori.

Varieties of Descriptive Two-Dimensionalism

That is the basic idea behind the two-dimensionalist revival of descriptivism. I will now outline several different versions of two-dimensionalism in more detail. I begin with *benign two-dimensionalism*, which is the view that there are two dimensions of meaning – character and content. Character is a function from contexts of utterance to content, which in turn determines a function from circumstances of evaluation to extensions. Characters, which are sometimes called *two-dimensional intensions*, are, as David Kaplan taught us, crucial to the semantics of context-sensitive expressions. It is Kaplan who gave us benign two-dimensionalism. In his sense, "we are all two-dimensionalists now." In recent years, however, the term *two-dimensionalism* has come to stand for something more ambitious – a cluster of views that attempt to use something like the distinction between content and character to explain, or explain away, all instances of the necessary aposteriori and the contingent apriori. I will sketch four versions of ambitious two-dimensionalism – the pragmatic version of Robert Stalnaker's 1978 paper, "Assertion," the strong semantic version suggested in the mid nineties by Frank Jackson's From Metaphysics to Ethics and David Chalmers'

The Conscious Mind, a weak semantic version which is a natural retreat from strong two-dimensionalism, and a hybrid version suggested by Chalmers in 2002.[13]

Stalnaker's Pragmatic Two-Dimensionalism

In "Assertion," Stalnaker accepted that Kripke's examples of the necessary aposteriori express necessary truths that predicate essential properties of objects. He also recognized that a speaker who assertively utters one of them asserts something that is knowable only aposteriori. However, he maintained that in every such case the proposition asserted is contingent, and so not identical with the proposition semantically expressed by the sentence uttered. He believed he could show this by appealing to a pragmatic model of discourse. According to the model, conversations take place against a background of shared assumptions which rule out certain possible world-states as not obtaining, or "being actual." As the conversation proceeds, new assertions acquire the status of shared assumptions, and the set of world-states compatible with what has been assumed or established shrinks. The aim of further discourse is to continue to shrink this set (called *the context set*) within which the actual state of the world is assumed to be located. The function of assertion is to eliminate from the context set all world-states in which the proposition asserted isn't true.

Stalnaker postulates three rules governing assertion.

R1.   A proposition asserted should always be true in some but not all members of the context set.

---

[13] Robert Stalnaker, "Assertion," Syntax and Semantics 9, 1978, reprinted in Stalnaker, Context and Content, (Oxford: Oxford University Press), 1999; Frank Jackson, From Metaphysics to Ethics, (Oxford: Clarendon Press), 1998; David Chalmers, The Conscious Mind, (New York and Oxford: Oxford University Press), 1996; and David Chalmers, "Components of Content," in David Chalmers, ed., The Philosophy of Mind: Classical and Contemporary Readings (New York and Oxford: Oxford University Press), 2002.

R2.   Any assertive utterance should express a proposition relative to each world-state in the context set, and that proposition should have a truth value in each such state.

R3.   The same proposition should be expressed relative to each world-state in the context set.

The rationale for R1 is that a proposition true in all world-states of the context set, or false in all such states, fails to perform the function of assertion – namely, to circumscribe the range of possibilities within which the actual world-state is located. Of course, this rule, like the others, allows for some flexibility in how it applies. If someone seems to say something that violates it, one may sometimes conclude that no violation has really taken place because the context set isn't

quite what one originally thought, or because the speaker didn't really assert, or mean, what he at first seemed to assert or mean. In such a case, a speaker may say something the literal interpretation of which would violate the rule, knowing full well that he will be reinterpreted in a way that conforms to it.

Stalnaker's rationale for R2 is that if an utterance violates it, then for some world-state w, the utterance won't determine whether w should remain in the context set or not. In explaining the rational for R3, he employs his notion of *the propositional concept associated with an assertion*, which is related to Kaplan's notion of the character of a sentence. For Stalnaker, a propositional concept is a function from world-states, considered as possible contexts of utterance, to propositions -- where propositions are nothing more than assignments of truth values to such states. The propositional concept associated with an utterance of S is a function that maps each world-state w of the context set onto such an assignment. This assignment can be thought of, roughly, as the proposition that would be expressed by S, if the context of utterance were to turn out to be w.

D	i	j	k
i	T	T	T
j	F	F	T
k	F	T	T

Example D represents the propositional concept associated with a use of a sentence at a moment in which the context set consists of the world-states i, j, and k. D tells us that if i is the state the world is actually in, then the proposition expressed by the speaker's utterance is the proposition that assigns truth to every world-state of the context set, if j is the actual world-state, then the proposition expressed assigns truth to k and falsity to i and j, and if k obtains, the proposition expressed assigns falsity to i and truth to the other two. Which proposition is expressed depends on which world-state actually obtains. Since the conversational participants haven't agreed on this, they won't know which proposition the speaker is asserting, and so they will be at a loss as to how to update the context set. In this sort of case, we have a violation of R3.

With these rules in place, Stalnaker is ready to explain assertive utterances of Kripkean examples of the necessary aposteriori. The key involves apparent violations of R3. He gives an example of hearing a woman speaking in the next room. I tell you, *That is either Ruth Marcus or Judy Thomson.* Since demonstrative pronouns and proper names are rigid designators, this sentence expresses either a necessary truth or a necessary falsehood, depending on who is in the next room. Let i be a world-state in which the woman speaking is Ruth, j be a state in which it is Judy, and k be a state in which it is Mrs. Clinton. The propositional concept associated with this utterance is then E.

E	i	j	k
i	T	T	T
j	T	T	T
k	F	F	F

E tells us two things: (i) we don't know which proposition is actually expressed by the sentence uttered, because we don't know which context actually obtains; (ii) none of the possible

propositions expressed would serve a purpose in the conversation. Thus, E violates R3, and any attempt to avoid violation by excluding i, j, or k from the context set would violate R1.   So, if we are to avoid violation entirely, and to regard the speaker's utterance as informative, we must take it as asserting something else. What else? Well, whichever world-state turns out to obtain, the speaker will be committed to his utterance expressing a truth. Thus, we take the proposition asserted to be one that is true at any world-state w (of the context set) just in case the proposition expressed by the sentence in w is true at w. This is the assignment of truth values that arises by looking along the diagonal in E to find the value that appears in row w of column w, for each w. Stalnaker calls this the *diagonal proposition.*   Since it is neither true in all world-states of the context set nor false in all those states, it can do what asserted propositions are supposed to do. Hence, it is what is really asserted by the speaker's utterance – no matter which member of the context set actually obtains.   This is what $\Uparrow$E represents, where '$\Uparrow$' is an operator that maps a propositional concept C1 onto the propositional concept C2 that assigns each potential context the diagonal proposition of C1.

$\Uparrow$**E**	i	j	k
i	T	T	F
j	T	T	F
k	T	T	F

This example is the prototype for Stalnaker's treatment of the necessary aposteriori. Since the example is supposed to generalize to all such instances, one might get the impression that he thought that all genuine cases of the necessary aposteriori involve indexical sentences, which semantically express different propositions in different contexts.   Since many examples of the necessary aposteriori involve only names or natural kind terms, such a view would require analyzing these terms as descriptions rigidified using either the actuality operator, or David Kaplan's *dthat* operator. However, Stalnaker didn't accept any such analysis. This created a problem for his pragmatic model. In order to get the desired results, he needed propositional

concepts associated with instances of the necessary aposteriori to assign different propositions

to different contexts. However, if the sentences aren't indexical, the needed propositional

concepts can't be their characters, since these will be constant functions. But then, it is not

evident where the needed propositional concepts will come from, or what interpretation they

should be given once we have them. Although Stalnaker tried to finesse this issue, it was a

problem.[14] Thus, it is not surprising that later two-dimensionalists took the step of analyzing

names and natural kind terms as indexical, rigidified descriptions – thereby transforming the

model from a pragmatic to a semantic one.

The standard argument for taking this step goes like this: Imagine finding out that

chemists have been wrong about the stuff the falls from the sky as rain, and fills the lakes and

rivers. It is not really H_2O but XYZ. What would *water* refer to, if this scenario were to turn out

to be actual? XYZ, of course. Although in the world as it really is, we use *water* to rigidly refer

to H_2O, in a possible context in which the scenario described is actual, *water* rigidly refers to

XYZ. Since the reference of *water* varies in this way from context to context, even though its

meaning remains the same, the ambitious two-dimensionalist concludes that it is indexical. Thus,

there must be some description implicitly associated with it by speakers that determines its

reference in different contexts, which is then rigidified. It may be difficult to determine precisely

what this description is, but that there is descriptive content to be rigidified is beyond question.

Ditto for every name and natural kind term.[15]

Strong and Weak Two-Dimensionalism

With this we move from Stalnaker's pragmatic model to contemporary semantic versions

---

[14] See my "Understanding Assertion" for a discussion of this point.

[15] David Chalmers's use of this line of reasoning is discussed at length in my <u>Reference and Description</u>, pp. 209-228.

ambitious two-dimensionalism. I begin with two versions, which I will call *strong* and *weak*.

Central Tenets of Strong Two-Dimensionalism

ST1. Each sentence S is semantically associated with a primary intension and a secondary intension. Its primary intension is a proposition which is true with respect to all and only those contexts C to which the character of S assigns a proposition that is true at C. The secondary intension of S at C is the proposition assigned by the character of S to C.

ST2. Understanding S consists in knowing its character and primary intension. Although this knowledge, plus relevant knowledge of a context C, would give one knowledge of the secondary intension of – i.e. the proposition expressed by -- S in C, one does not always have such knowledge of C. However, this does not prevent one from using S correctly in C.

ST3a. All names and natural kind terms have their reference semantically fixed by descriptive properties which can, in principle, be expressed by descriptions not containing any (ineliminable) names or natural kind terms.

ST3b. These terms are synonymous with descriptions rigidified using *dthat* or *actually*.

ST4a. ***It is a necessary truth that S*** is true with respect to a context C and world-state w iff the secondary intension of S in C is true with respect to all (metaphysically possible) world-states w* that are possible relative to w. Similarly for other modal operators.

ST4b.   ***It is knowable apriori that S*** is true w.r.t. C and w iff the primary intension of S in C is knowable apriori in w; ***x knows / believes (apriori) that S*** is true of an individual i w.r.t. C and w iff in w, i knows / believes (apriori) the primary intension of S in C.

ST5a. S is an example of the necessary aposteriori iff the secondary intension of S (in C) is necessary, but the primary intension of S is contingent and knowable only aposteriori. Such a sentence expresses a necessary truth in our actual context, while expressing falsehoods in other contexts. Its primary intension is not knowable apriori because we require empirical information to determine that our context is not one to which the character assigns a falsehood.

ST5b. S is an example of the contingent apriori iff the secondary intension of S (in C) is true, but not necessary, while the primary intension of S is necessary and knowable apriori.

Such a sentence expresses a proposition which is false at some world-states, even though it expresses a truth in every context. The primary intension of such a sentence is knowable apriori because no empirical information is needed to determine that its character assigns one's context a truth.

These theses carry with them certain more or less inevitable corollaries, including ST6a.

ST6a. There is no proposition that is both necessary and knowable only aposteriori; nor is there any proposition that is contingent yet knowable apriori.

If there were such propositions, then it should be possible to express them using nonindexical sentences, the primary and secondary intensions of which are identical (or at any rate equivalent). Since this is ruled out by ST5, the strong two-dimensionalist has reason to accept ST6a. A similar point holds for the corollary ST6b.

ST6b. The necessary aposteriori and the contingent apriori are, in effect, linguistic illusions born of a failure to notice the different roles played by primary and secondary intensions in modal and epistemic sentences.

In what follows, I will take strong two-dimensionalism to include ST6a and ST6b.

One final thesis is ST7, which adds the claim that all necessary truths are knowable.

ST7. A proposition is necessary iff it is knowable apriori.

In systems that identify propositions with sets of possible world-states, ST7 is trivial, and already presupposed. In fact, the acceptance by a strong two-dimensionalist of ST7 would seem to go hand in hand with a possible-world-state analysis of propositions. Since it is difficult to imagine another conception of propositions that would justify the claim that all necessary truths are knowable, it is difficult to understand why a strong two-dimensionalist would adopt ST7, unless the theorist wished to adopt that analysis of propositions. Why would such a theorist find such an analysis congenial? Well, consider the standard strong two-dimensionalist explanation of the contingent apriori, which hinges on the necessity of the sentence's primary intension. How does the necessity of this proposition guarantee that it is knowable apriori, unless it is guaranteed that every necessary truth is

knowable? And how is this guaranteed unless propositions are just sets of possible world-states?

To the extent that two-dimensionalists want simply to take it for granted that necessity of primary

intension is enough for apriori truth, they have reason to adopt the possible-world-state analysis of

propositions, and with it ST7. In what follows, I will call systems incorporating ST1–ST6 *strong*

two-dimensionalist, and those that also include ST7 *very strong two-dimensionalist*. I will take these

latter systems to identify propositions with sets of possible world-states.

Before giving a précis of weak two-dimensionalism, I pause over a point of

interpretation. Although the analysis of attitude ascriptions given in ST4 is crucial to strong

two-

dimensionalism, my best examples of strong two-dimensionalism – David Chalmers' The

Conscious Mind and Frank Jackson's From Ethics to Metaphysics – do not explicitly include ST4b,

or any other, semantic analysis of attitude ascriptions. Worse, Chalmers repudiates ST4b in

"Components of Content," published several years later.   Why, then, do I interpret the positions

taken in those books as suggesting strong two-dimensionalism? Because they strongly suggest

ST4, and because it is difficult without ST4 to maintain other theses to which Chalmers and

Jackson are committed. Detailed textual criticism aside, the main interpretive points are these:

(i)   Chalmers and Jackson hold that S is an instance of the necessary aposteriori iff the

          primary intension of S is contingent (and hence aposteriori) while the secondary

          intension of S is necessary, and that S is an instance of the contingent apriori iff the

          primary intension of S is necessary (and hence apriori) while the secondary intension

          of S is contingent. (These claims are themselves treated as necessary.)

 (ii)   Whenever S is an instance of the necessary aposteriori they endorse *It is not knowable*

          apriori that S. Whenever S is an instance of the contingent apriori, they endorse *It is*

          knowable apriori that S. In general, they endorse *It is knowable apriori that S iff the*

21

primary intension of 'S' is necessary, and they take it that (for any context C) the left-hand side expresses a truth (in C) iff the primary intension of S is necessary.

(iii)   From this it is reasonable conclude that, in their view (at the time), *it is knowable apriori that* operates on the primary intension of S, or if it operates on the primary intension of S plus something else, only the primary intension matters.

(iv)   There is no reasonable option to assuming that *know* and *believe* operate on whatever *it is knowable apriori* does.

In effect, this adds up to the analysis given in ST4b.[16]

Next, I will sketch weak two-dimensionalism. Theses WT1-WT4a differ only in minor matters of detail from ST1-ST4a, and for our purposes may be assumed to come to essentially the same thing. The key differences are in WT4b and WT5.[17]

Central Tenets of Weak Two-Dimensionalism

WT1. Each sentence S is semantically associated with a primary intension and a secondary intension. The former is its Kaplan-style character. The secondary intension of S at a context C is the proposition assigned by its primary intension to C.

WT2. Understanding S consists in knowing its primary intension. Although, this knowledge, plus relevant knowledge of the context C, would give one knowledge of the proposition expressed by S in C, one does not always have such knowledge of C. However, this does not stop one from using S correctly in C.

WT3. As before.
WT4a. As before.

WT4b. An ascription *x v's that S*, taken in a context C, is true of an individual A w.r.t. a world-state w iff there is some character M such that (i) in w, A bears R to M, and (ii) M assigns the secondary intension of S in C to a related context with A as agent and w as

---

[16] For detailed discussions of the two works see chapters 8 and 9 of <u>Reference and Description</u>.

[17] The differences between strong and weak two-dimensionalism are discussed in more detail in chapter 7 of <u>Reference and Description</u>.

world-state. So propositions are objects of the attitudes, and attitude verbs express two-place relations between agents and propositions. However, this two-place relation holds between A and p in virtue of a three-place relation holding between A, a character, and p. To believe p is to accept a character M that expresses p (and believe that M expresses a truth). To know p is to justifiably accept a character M that expresses p (and know that M expresses a truth).

WT5a   For all necessary propositions p, p is knowable only aposteriori iff (i) p is knowable by virtue of justifiably accepting some meaning M (and knowing that M expresses a truth) -- where (a) M assigns p to one's context, (b) M assigns a false proposition to some other context, and (c) one's justification for accepting M (and believing M to express a truth), requires one to possess empirical evidence -- and (ii), p is knowable only in this way.

WT5b   For all contingent propositions p, p is knowable apriori iff p is knowable by virtue of justifiably accepting some meaning M (and knowing that M expresses a truth) -- where (a) M assigns p to one's context, (b) M assigns a truth to every context, and (c) one may be justified in accepting M (and believing M to express a truth) without empirical evidence.

Although some of the differences between strong and weak two-dimensionalism are subtle and far-reaching, others are obvious. For example, whereas the strong and the weak two-dimensionalist agree that necessary-aposteriori sentences are always associated with two propositions – one necessary and one contingent – the weak two-dimensionalist maintains that the necessary proposition is itself knowable only aposteriori. Because of this, the weak two-dimensionalist cannot identify propositions with sets of possible world-states. Thus, although strong and weak two-dimensionalists both presuppose that whenever a primary intension is necessary, it is knowable apriori, the weak two-dimensionalist does not have the ready explanation of why this should be so that the very strong two-dimensionalist has. For the latter, it is in the nature of the one necessary proposition that it should be knowable apriori. Since the weak two-dimensionalist cannot say this, a different explanation is needed – of what sort is not obvious. This is one way in which weak two-dimensionalism is initially less attractive than

strong two-dimensionalism. However, this initial disadvantage is not decisive, since all three

versions of ambitious two-dimensionalism face crippling problems which require their

rejection. I will sketch a few of these problems, before turning to the final version of two-

dimensionalism to be considered.

Problems with Ambitious Two-Dimensionalism

Critique of Stalnaker's Pragmatic Two-Dimensionalism

      Here is one problem for Stalnaker's pragmatic account of the necessary aposteriori. I

hold up my briefcase. You look at it closely, and ask *What is it made of? Is it cow leather,*

some other kind of leather, vinyl, or what? I answer, *It is made of cow leather.* Let's assume

that although you don't know, prior to my utterance, what the briefcase is made of, we both take

it for granted that, whatever it is made of, it is an essential property of this briefcase that it be

made of that stuff. Since it is, in fact, made of cow leather, my remark is an example of the

necessary aposteriori. How would Stalnaker represent the conversation? Well, prior to the

utterance he would have different possible worlds-states in the context set that are compatible

with everything we assumed or established up to then. Presumably, these would include a

context i in which the one and only briefcase I am holding is made of cow's leather, a context j

in which the briefcase I am holding is made of something else, say, pigskin, and context k in

which I am holding a briefcase made of something else again – vinyl. So, he would associate

my remark with the propositional concept B.

B	i	j	k
i	T	T	T
j	F	F	F
k	F	F	F

      His rules for assertion would then yield two conclusions: (i) that on hearing my utterance

you had no way of knowing which proposition was expressed, because you didn't know which

context – i, j, or k – actually obtained, and (ii) that none of the propositions assigned by B to

these world-states would have served a useful purpose. To have asserted a necessary truth would have been uninformative, and to have asserted a necessary falsehood would have been a nonstarter. So, if you were to regard my utterance as informative, you had to take it as asserting some proposition other than any of the candidates assigned to members of the context set by B. Since you knew that whatever the real context turned out to be, I would be committed to my remark expressing a truth, you took me to have asserted the diagonal proposition, which is true at any of i, j, or k iff the proposition B assigns to that world-state is true when evaluated at that state. Since this proposition is neither true at all the world-states, nor false at them all, asserting it does the job that assertion is intended to do.

That is Stalnaker's account. There are two things wrong with it. First, it is wrong to suppose that you had any relevant doubt about what proposition was expressed by my utterance of *It is made of cow leather* . The proposition I expressed is one that predicates being made of cow leather of one particular briefcase – the one I was holding. You knew it was the object you had asked about, and about which I gave my answer. Since you also knew what cow leather was, you knew precisely which property was predicated of which object by my remark. How, then, could you have been in any real doubt about which proposition my sentence expressed?

The second thing wrong with the explanation is that world-states j and k in the context set must either be ones that are not really metaphysically possible (contrary to the assumptions of the model), or ones that are not compatible with all the shared assumptions prior to my utterance (also contrary to the model). What are the world-states i, j, and k? They are total possibilities regarding how the world might be in which I am holding one and only one briefcase, which is both seen by us and the subject of our discourse. The briefcase satisfying these conditions in i is made out of cow leather, whereas the briefcases satisfying them in j and k are made out of pigskin and vinyl, respectively. Which briefcases are these in j and k? If j and k really are metaphysically possible, as Stalnaker insists, then the briefcases there can't be

the briefcase I was really holding when I answered your question. Since that briefcase is made out of cow leather in every genuinely possible world-state in which it exists, it is not made out of pigskin in j, or vinyl in k.   It follows that j and k must be world-states in which I am holding some other briefcase. But how can that be? Surely, one thing we both knew prior to my remark was that I was holding this very briefcase [imagine me holding it up again now], which we saw and were talking about. But if we did have this *de re* knowledge, or these *de re* beliefs, then Stalnaker's requirement that the world-states in the context set be compatible with everything assumed and established in the conversation must have eliminated all metaphysically possible world-states in which other briefcases were under discussion. But if that is right, then there is no room for the diagonalization required by his explanation, and the account fails.

This example relies on a plausible, but potentially contentious, metaphysical doctrine -- the essentiality of origin, or constitution. However, there is nothing special about the particular essential property chosen. Other essential properties or relations (e.g. the property of being non-identical with Saul Kripke, or the relation of non-identity itself) would serve equally well. The important thing is simply that there be such properties (and relations). Given that there are, we can reconstruct different versions of the same problem that don't rely on assumptions about material constitution. For example, if, pointing at a man, David Kaplan, whom we both clearly see and know we have been talking about, you ask *Is he Saul Kripke?* and I reply *No, he isn't Saul Kripke*, what I say is knowable only aposteriori, even though the proposition expressed by my sentence is a necessary truth. However, if we try to apply Stalnaker's two-dimensionalist model of discourse to this example, we will run into precisely the same problems we did in the example about my briefcase. Instances of the necessary aposteriori that hinge on the essential properties of things cannot, in general, be explained by the model.

Notice what would happen if we dropped one of the antecedent philosophical commitments used in constructing the model – the restriction of the epistemically possible to the metaphysically possible. The idea is to allow the context set to include world-states that are metaphysically impossible, but epistemically possible – i.e. maximally complete properties that the world couldn't really have had, but which we cannot know apriori that it doesn't have (on analogy with certain properties, like being made of vinyl, that my briefcase couldn't really have had, but which one can't know apriori that it doesn't have). When we allow such world-states, the propositional concept associated with the utterance turns out to be different from what we originally took it to be. On this way of looking at things, i, j, and k are different epistemic possibilities involving the very same object – the briefcase I was actually holding. In i, it is made of cow leather; in j, it of pigskin; and in k, of vinyl. This gives us the propositional concept B*.

B*	i	j	k
i	T	F	F
j	T	F	F
k	T	F	F

Since the same proposition is expressed with respect to each possible context, and since that proposition is neither trivially true nor trivially false, diagonalization is irrelevant.[18]

<u>Critique of Strong Two-Dimensionalism</u>

The most obvious problem for strong two-dimensionalism is that its account of the semantics of attitude ascriptions is obviously incorrect. When ordinary indexicals like *I, now,* and *he* are involved, their secondary intensions in contexts are crucial for attitude ascriptions in which they occur in the complement clause. Since the thesis ST4b doesn't make room for

---

[18] Stalnaker's two-dimensionalist account of assertion, the objections to it, and the distinction between aspects of it that must be abandoned and those that can be retained are discussed at greater length in my "Understanding Assertion."

this, it clearly fails. However, strong two-dimensionalism can be shown to be false, even if these indexicals are put aside.

Here is one argument. Let n be a name. If strong two-dimensionalism were correct, then for some F, n would be synonymous with **the actual F,** and hence would have the same primary intension as **the F** – in which case, the primary intensions of **n is G** and **the F is G** would be the same. If , in addition, attitude ascriptions reported relations to the primary intensions of their complement clauses, then (5a) and (5b) would have the same secondary (as well as primary) intensions, and so would be necessarily equivalent. This would mean that (6a) and (6b) couldn't differ in truth value.

5a.   A believes (knows) that n is G
 b.   A believes (knows) that the F is G

6a.   If it had been the case that …, then it would (or would not) have been the case that … A
        believed (knew) that n was G ….]
 b.   If it had been the case that …, then it would (or would not) have been the case that … A
        believed (knew) that the F was G ….]

 Of course, pairs of the form (6a) and (6b) can easily differ in truth value, as is shown by (6aa,6bb).

6aa.   Although John truly believes (knows) that n is F, had the world been in state w, n would
        not have been F and John would not have believed (known) that n was F.
6bb.   Although John truly believes that the F is F, had the world been in state w, the actual F
        would not have been F and John would not have believed (known) that the F was F.

The lesson here is that since n is rigid, its referent o is such that, at any world-state w, **n is F** is true iff at w it is a fact that o has the property expressed by F; thus, it ought to be the case that, for any world-state w, **John's belief that n is F** stands for a belief about o – one that is true at w only if, at w, o has the property expressed by F. Incredibly, on strong two-dimensionalism this turns out not to be so; instead "the fact that n is F" and "the belief that n is F" are, wrongly, allowed to be about different individuals.

28

Here is another argument that makes a similar point. Where n is a name that refers to o, for any world-state w, if ***John believes (knows) that n is G*** is true with respect to w, then in w John believes (knows) of o that it "is G". In other words, when n designates o, (7b) must be true, if (7a) is.

7a. In world-state w [A believes (knows) that n is G]
  b. In world state-w [A believes (knows) that x is G] (relative to an assignment of the referent of n to 'x')

Strong two-dimensionalism can't capture this, since it wrongly allows the ascription ***A believes (knows) that n is G*** to be about different individuals with respect to different world-states. For reasons like these, it is clear that strong two-dimensionalism is false.[19]

Critique of Weak Two-Dimensionalism

Although the issues here are somewhat more complex, in the interest of economy I will be brief.[20] There are two main points to make. First, the arguments that the reference of names and natural kind terms must be semantically fixed by description are defective, and do nothing to rebut the original Kripkean arguments that the vast majority of such terms do not have their referents fixed in this way. Second, even if we had reference-fixing descriptions, there would be no way of rigidifying them that would meet the needs of the weak two-dimensionalist.

Regarding the first point, about the lack of reference-fixing descriptions, one must distinguish (a) foundational facts that bring it about that terms acquire (and retain) the semantic properties they have from (b) semantic facts about them, known by competent speakers, that constrain what they refer to -- in the way that being constrained to refer to a female is part of what must be known in order to understand the demonstrative *she*. When it is claimed by Jackson, Chalmers, and others that the reference of names and natural kind terms must be

---

[19] For additional arguments, see chapter 10 of <u>Reference and Description</u>.

[20] For a more complete discussion see chapter 10 of <u>Reference and Description</u>.

fixed by description, they are mixing up these two things. For example, when Jackson

claims that there must be reference-fixing descriptions associated with names and natural kind

terms, because, as he puts it in <u>From Metaphysics to Ethics</u>, [21] "it isn't magic" that they refer to

what they do, he is pointing to the fact that these expressions – like all others – get to have the

semantic properties they do in virtue of some describable empirical features of how they are

used. Of course, they do; and, of course, being language users, we are not entirely ignorant of

what these features are. However, it doesn't follow that we have complete and accurate

knowledge of them. Still less does it follow that such knowledge is part of what a speaker

must understand in order to understand the expression. To understand a word you have to

know what it means -- which in certain cases may amount to knowing its reference -- but you

don't have to know how the word got to have the meaning or reference it does. So a guarantee

that there is some accurate description of how a word got to have its meaning and reference

does not provide any guarantee that the word has its reference semantically fixed by a

description grasp of which is required for linguistic competence.

      This point bears on the suggestion, frequently made by contemporary descriptivists,

that what Kripke did in giving his historical-chain account of reference transmission was

simply to offer his own version of the description theory of reference-fixing.[22] The idea

behind this suggestion is, in simplest terms, that the reference of a name n for a particular

speaker is determined by something like the following the description: *the individual which*

the person or persons from whom I acquired n referred to when they used the name. One of

the things that makes this idea seem plausible is the requirement, recognized by Kripke, that

in order for a reference-determining chain to be created by passing a name from speaker to

---

[21] Page 82.

[22] See, for example, the passage from Lewis referenced in fn. 11.

speaker, the person acquiring the name must intend his reference to be parasitic on the reference of his source(s). The descriptivist proposes simply to state this requirement in terms of a reference-fixing description.

Although the attraction of this idea is evident, when one examines it more closely one finds ample reason to be skeptical. It is not clear that speakers invariably have in mind, among all the different descriptions they associate with a given name, some accurate and precise reference-fixing description for it. Certainly, the description *the referent of the name as it was used by the first person from whom I acquired it* does not always pick out my present referent for it.[23] Since theorists have yet to develop a complete, accurate, and fully explicit historical-chain theory from which the needed description could be extracted, it is not clear which, if any, such description would be sufficient to handle all problematic cases. Even if theorists were to come up with the desired theory, it is far from obvious that ordinary speakers must have had it at their disposal all along, whenever they used a proper name.

We may assume that there is some process of reference-borrowing by which later uses of a name inherit their reference from earlier uses. We may also assume that speakers know that some such process exists, even though they don't know precisely how it works. With this in mind, consider the parasitic, reference-borrowing description D.

D.   the thing referred to by those uses -- whatever they turn out to be -- of the name n
        from which my present use of n inherits its reference.

_____

[23] Suppose, for example, that the person from whom I first picked up the name *Plato* was talking about his neighbor, whom he believed to be very wise. Suppose further that after speaking to this person I had many other conversations in which the name was used to describe Socrates' famous biographer, whom I later started to read about under the name *Plato*. All of this could be true, even if I wrongly assumed that the person from whom I first

Since it is plausible to suppose that speakers know, at some level, that the reference of their uses of names is inherited from the use of these names by other speakers – even though they don't know the precise mechanism by which this takes place – it is reasonable to suppose that they implicitly associate D, or something very much like it, with each name n that they use. However, it is important not to be confused by this. Although D does denote the referent of the speaker's use of n, the reference of n is not semantically fixed by D. If it were, then satisfaction of D would be the mechanism by which the speaker's use of n acquired its reference. But this can't be the way in which reference is determined, since D itself presupposes that the speaker's use of n already has a reference, which has been inherited from something else. In order for D to correctly pick out the referent, the speaker's use of n must have acquired that referent in some other way. Thus, the mechanism by which its reference is fixed can't be via satisfaction of D.

The upshot of all this is that even though the reference-borrowing facts of Kripke's historical-chain account of reference transmission play a foundational role in determining the reference of names and natural kind terms for speakers, there is no reason to think that these facts are among the semantic facts that must be mastered by competent speakers. Since this is precisely what would be required if the referents of these terms were semantically fixed by description, the description theory of reference fixing remains unsupported.

Similar considerations can be used to scotch a familiar two-dimensionalist objection to Kripke's semantic arguments against the description theory of reference fixing. These arguments are based on thought experiments which show that often, in different counterfactual circumstances of use, a name or natural kind term n refers to something other than what is denoted by the descriptions that speakers explicitly associate with n, and that they would offer in

---

heard the name was talking about the same individual as everyone else. In this sort of case, I do not refer to the neighbor of my original source when I use the name *Plato*. Instead, I refer to the famous philosopher.

answer to the question ***To whom, or what, do you use 'n' to refer?*** From this, Kripke

concludes that n does not have its reference fixed by descriptions associated with it by speakers.

On the contrary, the two-dimensionalist objects, Kripke's thought experiments undermine his

conclusion by covertly presupposing unarticulated reference-fixing descriptions that implicitly

guide speakers' judgments.

Here is Frank Jackson.

> Our ability to answer questions about what various words refer to in various
>
> possible worlds, it should be emphasized, is common ground with critics of the
>
> description theory. The critics' writings are full of descriptions (*descriptions*)
>
> of possible worlds and claims about what refers, or fails to refer, to what in
>
> these possible worlds. Indeed, their impact has derived precisely from the
>
> intuitive plausibility of many of their claims about what refers, or fails to refer,
>
> to what in various possible worlds. But if speakers can say what refers to what
>
> when various possible worlds are described to them, description theorists can
>
> identify the property associated in their minds with, for example, the word
>
> 'water': it is the disjunction of the properties that guide the speakers in each
>
> particular possible world when they say which stuff, if any, in each world
>
> counts as water. This disjunction is in their minds in the sense that they can
>
> deliver the answer for each possible world when it is described in sufficient
>
> detail, but it is implicit in the sense that the pattern that brings the various
>
> disjuncts together as part of the, possibly highly complex, disjunction may be
>
> one they cannot state.[24]

---

[24] "Reference and Descriptions Revisited," p.212.

Jackson maintains that our ability to identify the reference of terms in imagined

Kripkean scenarios presupposes a knowledge of complete and accurate reference-fixing

descriptions that is adequate for every scenario. He takes our ability to make these

Kripkean judgments to demonstrate that names and natural kind terms must have their

reference semantically fixed by these descriptions, that competence with the terms requires

mastery of the descriptions, and that the terms themselves have the semantics of (indexical)

rigidified descriptions. However, the ability he points to – such as it is – is not sufficient to

draw these conclusions.

      First, there are clear cases in which we have no trouble identifying the referent of a

term t, even though it is clear that there is no reference-fixing description associated with t

by speakers. David Kaplan's example of the identical twins, Castor and Pollux, discussed

earlier, is a case in point. We have no trouble identifying Castor as the referent of his use of

I, and Pollux as the referent of his, just as we have trouble recognizing ourselves as referents

of our own uses.   This is so despite the fact that the referent of *I* is not semantically fixed,

for any of us, by descriptions we semantically associate with it. If this is true of *I*, it is

surely also true of *now*, and may be true of other expressions as well. Second, even in cases

in which there may be descriptions picking out the referent of a term that are, in some sense,

associated with it by speakers, it remains to be shown these descriptions play any role in its

semantics. One can describe possible scenarios in which our intuitions tell us that speakers

use the word *and* to mean disjunction, the material conditional, the property of being a

necessary truth, or the property of being a philosopher. Even if one were to grant the

assumption that these intuitions arise from an internalized theory T that unconsciously

guides us, it would not follow that the meaning of *and* – it's character in Kaplan's sense – is

one that yields as content in a context whatever satisfies the relevant description extractable

from T. Surely not every word is a descriptive indexical in Kaplan's sense.   To miss this

point is to miss the distinction noted earlier between (a) foundational facts that bring it about that terms acquire (and retain) their semantic properties from (b) semantic facts about them, known by competent speakers, that constrain what they refer to. Whereas the descriptivist needs reference-fixing descriptions arising from (b), Jackson's argument can't exclude the possibility that the only relevant descriptions are those arising from (a). Finally, the claim that our ability to categorize cases in certain ways presupposes the sort of underlying knowledge required by the description theory is tendentious in something like the way that Plato's attribution of a priori knowledge of mathematics to the slave boy in the *Meno* is tendentious. There are other ways to explain the recognition of new facts.

The failure of Jackson's argument on this point is representative of the arguments of two-dimensionalists generally on the subject of descriptive reference fixing. Typically these arguments do not involve precise and detailed proposals about the descriptions that allegedly fix the referents of specific names or natural kind term. Instead, they attempt to establish, on the basis of very general considerations applying to all such terms, that their reference simply must be fixed descriptively – even if we are not in a position to articulate the crucial descriptions themselves.   As the discussion of Jackson has illustrated, these arguments fail for a number of reasons, most notably the failure to distinguish foundational from semantic facts. Once this is recognized, and the Jackson-style arguments are out of the way, full weight can be given both to Kripke's refutations of particular proposals for treating specific descriptions as semantically fixing the reference of particular terms, and to the fact that complete, accurate, and reliable descriptions, known by speakers to fix the reference of typical names and natural kind terms, have not been formulated by anyone.

The next point to be made is that even if reference-fixing descriptions were generally available for names and natural kind terms, the weak two-dimensionalist has no good way of rigidifying them. One of two standard ways of rigidifying a description is by adding the

actuality operator. However, this is no help to the weak two-dimensionalist, since it is

obvious that neither names nor natural kind terms are equivalent to descriptions rigidified

using this operator. For example, *Aristotle* isn't equivalent to **the actual F**, for any F – since it

is possible for agents in other possible world-states to know or believe that Aristotle was a

genius without knowing or believing anything about our actual world-state -- whereas it is not

possible for those agents to know or believe that the actual so and so was a genius without

knowing or believing anything about our actual world-state.   In the presence of the defining

assumptions of weak two-dimensionalism, this means that the propositions expressed by (i.e.

the secondary intensions of) **Aristotle was a genius** and **The actual F was a genius** cannot be

the same, no matter what F one chooses.[25]

Thus, if the weak two-dimensionalist is going to take names and kind terms as rigidified

descriptions, he will have to take them to be descriptions rigidified using David Kaplan's *dthat*

operator.[26] However, this choice is also problematic. For one thing, it makes weak two-

dimensionalists Millians about semantic contents – and hence subject to the very problems

posed by Frege's puzzle and Russell's problem of negative existentials that descriptivists

standardly take to refute Millianism, and to motivate descriptivism in the first place.   In

addition, the weak two-dimensionalist explanations of the necessary aposteriori and the

contingent apriori go disastrously awry if one holds (in accord with WT4b) that understanding

and justifiably accepting **dthat [the F] is G** is sufficient for knowing the (Russellian)

---

[25] This argument is elaborated in chapter 2 of my <u>Beyond Rigidity</u>, (New York: Oxford University Press), 2002.
See also pp.303-6 of <u>Reference and Description</u>.

[26] Another reason for weak two-dimensionalists to forsake analyses involving the actuality operator is directly
connected to the point made in fn. 12. Given an instance, $n \neq m$, of the necessary aposteriori, one cannot analyze
the names n and m as *the x: actually Nx* and *the x: actually Mx*, unless *~the x: actually Nx = the x: actually Mx* is
knowable only aposteriori – something called into question by the argument mentioned there.

proposition it expresses. It is easy to show that, on this view, the necessary aposteriori will

shrink to the vanishing point, and the contingent apriori will bloat to include virtually all

contingent propositions involving individuals or kinds.[27] To block this *reductio ad absurdum*,

one must add a condition -- knowing *de re* of the denotation of the *dthat*-rigidified description

that it is so denoted – to what is required in order for understanding and justifiably accepting

dthat [the F] is G to count as knowing p, where p is the proposition it expresses. But then there

is a further difficulty. The *de re* knowledge mentioned in the condition is itself knowledge of a

singular proposition. How does it arise –for all names and natural kind terms for which *dthat*-

rigidified descriptions are posited as analyses by weak two-dimensionalists, i.e., for every name

and kind term? Since this knowledge can't typically be explained as the result of understanding

and (justifiably) accepting still further indexical sentences, some additional, non-two-

dimensionalist, explanation of the crucial *de re* knowledge and belief must be given. If we had

such an explanation, however, we could, presumably, apply it directly to paradigmatic Kripkean

examples in which we know *de re* of an object that it has an essential property, even though this

knowledge can only be aposteriori. But then we have an instance of the Kripkean necessary

aposteriori that can't be forced into the weak two-dimensionalist mold, and the weak two-

dimensionalist account of the necessary aposteriori is subject to a falsifying counterexample.[28]

A Hybrid View

Confronted with these objections to existing versions of ambitious two-dimensionalism,

one might wonder whether there is some other version that is immune to the problems we have

found. I don't think there is. However, the matter is not easily settled, since it is not clear what

modifications of two-dimensionalism are possible without abandoning essential features of the

---

[27] Reference and Description, pp. 307-310.

[28] Ibid. pp. 310-313.

program. I will say a word about this by considering a proposal put forward by David

Chalmers, in his 2002 paper, "The Components of Content," where he sets his sights on solving

Frege's puzzle by treating names and natural kind terms as rigidified descriptions, and making

attitude verbs sensitive to both primary and secondary intensions. The new approach implicitly

rejects the strong two-dimensionalist analysis of attitude ascriptions given by ST4b and the weak

two-dimensionalist analysis given by WT4b, in favor of something along the lines of H.

H.   Both the primary and secondary propositions (intensions) associated with S are responsible
        for necessary conditions on the truth of *x (knows) believes that S*, without providing
        sufficient conditions. Such an ascription, as used in C, is true of an agent A in a
        circumstance of evaluation w iff in w, A (justifiably) accepts some (true) sentence or
        mental representation M which is such that (i) the secondary intension of M in A's context
        in w is identical with the secondary intension of S in C, and (ii) the primary intension of M
        is "appropriately related" to the primary intension of S in C.

        There are various reasons to doubt the effectiveness of this proposal in solving Frege's

puzzle. For one thing, it presupposes that all names and natural kind terms have their referents

semantically fixed by description, and so is vulnerable to the first objection brought against weak

two-dimensionalism. For another thing, the most plausible candidates for reference-fixing

descriptions that might withstand Kripke's semantic arguments involving error, misdescription,

and failures of uniqueness are metalinguistic descriptions involving the historical chain of

reference transmission in which the user of the term stands. However, these descriptions are

very poor candidates for specifying the contents of the beliefs and assertions one ascribes to

others. When I say that the ancient Babylonians believed that Hesperus was a star, I am surely

not saying that their mental contents included anything having to do with the chain of reference

transmission the brought the name *Hesperus* to me. For these, and other, reasons Chalmers'

proposal for solving Frege's problem doesn't seem very promising.[29]

But even if we put Frege's puzzle aside, adopting H is a more effective strategy for

challenging ambitious two-dimensionalism than for saving it. Such a strategy threatens the

conjunction of two principles, Apriori 1 and Apriori 2, which have been central to the approach.

Apriori 1
A sentence S is an instance of the apriori iff the character of S assigns every context a
proposition that is true in that context.

Apriori 2
It is knowable apriori that S is true iff S is an instance of the apriori.

Since the objection against taking names to be descriptions rigidified using the actuality operator

within the framework of weak two-dimensionalism remains in force against the hybrid version of

the view, the hybridist must analyze names and natural kind terms as *dthat*-rigidified

descriptions. But then, by Apriori 1, *If there is a unique F, then n is F* is analyzed as *If there*

is a unique F, the dthat [the F] is F, and so is characterized as apriori – when the referent of n is

fixed by *the F*. The right-to-left direction of Apriori 2 will then characterize the ascriptions that

result from prefixing *it is knowable apriori that* to them as true. However, this can be correct

only if *It is knowable apriori that if there is a unique F, then x is F* is true relative to an

assignment of the referent of n to 'x'.[30] Since such *de re* knowledge of individuals can't, in

---

[29] Ibid., 322-324.

[30] It might seem that the defender of H could deny the move from the truth of *It is knowable apriori that if there is*

a unique F, then n (i.e. dthat [the F]) is F to the truth of *It is knowable apriori that if there is a unique F, then x*

is F relative to an assignment of the referent of n to 'x', on the grounds that the primary intension of the

complement clause of the latter is not "appropriately related" to that of the complement clause of the former.

However, this would fly in the face of the fact that if n is a name that refers to o, the truth of *α knows / believes*

general, be apriori, the right-to-left direction of Apriori 2 is in jeopardy.   Nor would denying

this be of much help, since if such knowledge of individuals can be apriori, then ***If there is a***

unique F, then x is F (taken relative to an assignment of the referent of n to 'x') will, according

to Apriori 2, count as an instance of the apriori, even though, typically, its character (relative to

such an assignment) will not assign a truth to every context -- thereby falsifying the conjunction

of the left-to-right directions of Apriori 1 and Apriori 2.

An independent argument against this conjunction can be constructed as follows.

Step 1: If H is correct, then there will be cases in which an utterance, by a_1 in context C_1, of
A_2 knows that n is F is true, where (i) the referent a_2 of A_2 justifiably accepts a true
sentence ***m is F*** the secondary intension of which in a_2's context C_2 is the same as
the secondary intension of ***n is F*** in C_1, and (ii) the primary intensions of these two
sentences are at least somewhat different because the primary intension of m (as
used in C_2) differs slightly from that of n (as used in C_1). Here, n is a name, and m
is either a name or an indexical -- m may even be the same name as n, provided that
one recognizes (as Chalmers knows he must) that different speakers may use it with
somewhat different primary intensions in their respective contexts.

Step 2: If there are cases of the sort indicated in Step 1, then we may suppose that some will
be symmetrical in that what a_1 uses ***A_2 knows that n is F*** to truly report about a_2, the
latter can use ***A_1 knows that m is F*** to truly report about a_1. These reports can be
jointly true because the secondary intensions of the two complement clauses are the
same in their respective contexts, while the primary intensions of these clauses,
though different, are close enough to satisfy constraints imposed by H.

Step 3: We may assume, with the two-dimensionalist, that the primary intensions of n and
m are given by a pair of descriptions D_n and D_m which pick out the same objects in
their respective contexts, but which denote different objects with respect to some

---

(apriori) that …n… guarantees the truth of ***α knows / believes (apriori) that …x…***, relative to an assignment of o

to 'x'. See pp. 261-262 and 316-318 of <u>Reference and Description</u>.

pairs of contexts. One way for this to happen is for D_n to be like D_m except for containing a conjunct that predicates G of its denotation whereas the corresponding conjunct in D_m predicates n unrelated predicate F.

Step 4: Now consider a_1's utterance in C_1 of the ascription *A₂ knows that if n exists, then n is F*. For the reasons given in Step 2, this ascription should be true, since a_2 justifiably accepts the true sentence *If m exists, then m is F*, the secondary intension of which in C_2 is the same as the secondary intension of *If n exists, then n is F* in C_1.   As before, although the primary intensions of the two sentences (as used in their respective contexts) differ slightly, they should be close enough to satisfy the constraints imposed by H.

Step 5: Since F is included in the reference-fixing conditions for m, the primary intension of *If m exists, then m is F* (as used in C_2) should be necessary, and so the knowledge correctly attributed to a_2 by a_1's utterance of *A₂ knows that if n exists, then n is F* should (according to ambitious two-dimensionalism) be counted as apriori.

Step 6: So *A₂ knows apriori that if n exists, then n is F* should be true, as used in context C_1, as should *It is knowable apriori that if n exists, then n is F* -- even though the primary intension of *If n exists, then n is F*(as used in C_1) is not necessary (thereby falsifying the conjunction of the left-to-right directions of Apriori 1 and Apriori 2).

Conclusion

Having reached this point, we may summarize our results as follows: The versions of ambitious two-dimensionalism we have been able to state precisely – pragmatic two-dimensionalism, plus strong and weak semantic two-dimensionalism – are clearly incorrect. [31] The hybrid version of the view – which is problematic as a solution to Frege's puzzle – itself threatens principles central to ambitious two-dimensionalist accounts of the necessary aposteriori

[31] For a different, but in some respects related, argument for a similar conclusion about the prospects for two-dimensionalism, see George Bealer's discussion in section 2 of "Modal Epistemology and the Rational Renaissance," in Tamar Gendler and John Hawthorne, eds., Conceivability and Possibility (Oxford: Oxford University Press), 2002.

and the contingent apriori. Since there is no coherent doctrine of ambitious two-dimensionalism without such accounts, there is, at present, no further version of the view left to consider. If the two-dimensionalist approach is to survive at all, the onus is on those who favor it to formulate a new version that avoids the objections raised here. For that, we will simply have to wait.