The End of Descriptivism

by

Sam Alexander McNutt
Class of 2013

A thesis submitted to the faculty of Wesleyan University in partial fulfillment of the requirements for the Degree of Bachelor of Arts with Departmental Honors in Philosophy

Middletown, Connecticut April, 2013
Acknowledgements

I would like to thank my advisor Sanford Shieh without whose help this thesis would not have been possible. I am grateful to everyone who has been willing to discuss these matters with me or even simply listen as I tried to put ideas into words. Any errors in this work should be attributed to me alone.
Chapter 1:
History and Implications of a Discussion about Language

Language is the medium of thought and our primary means of communicating abstract ideas. It is thus unavoidably central to philosophy. It is often assumed that linguistic expressions, at least those for which questions of truth are relevant, must somehow refer to something outside language. It is, however, not immediately clear how such reference is achieved. I will discuss various theories of how such reference takes place. Given the course recent discussion of this topic has taken, I will mostly confine my arguments to the reference of singular terms.

Philosophical discussion of language in the West goes back at least to Plato. In the *Cratylus*, the character of Socrates argues that names (*onomata*) are introduced into language as abbreviated descriptions of their referents. For example, he analyzes the word for man, ‘*anthrôpos,*’ as an abbreviation of ‘*anathrôn ha opôpe,*’ meaning ‘one who reflects on what he has seen.’ On Plato’s view, names implicitly describe their referents. While this analysis may seem compelling in the case of ‘*anthrôpos,*’ one doubts whether it would be so successful analyzing, say, ‘blimp.’

John Stuart Mill proposed a more superficially plausible account of naming. According to Mill, proper names have denotation but not connotation. That is to say, names have reference but lack descriptive content. Typically, Mill is interpreted as claiming that the meaning of a name is its referent and nothing more. It’s not entirely clear that this is actually Mill’s view, but I will assume this standard interpretation when referring to Mill and Millianism. For all its immediate appeal, Mill has no account of how exactly names manage to refer. If, when a speaker learns a name, no descriptive content is learned, it is difficult to imagine how the speaker could associate
the proper denotation with the name, at least in cases in which the denotation were not known to
the speaker demonstratively.

Failure to explain how names denote is not the only problem with Millianism. Various
linguistic conundrums suggest that the meaning of names is not exhausted by their denotation.
For instance, it is obvious that statements of identity involving co-referential names often convey
information, or have cognitive value. Were the meaning of names simply their denotation, such
statements could never be informative. They would be mere tautologies, expressions of some
object’s self-identity. This dilemma is often called “Frege’s Puzzle.” The meaningfulness of true
negative existential statements, i.e. statements of the form ‘X does not exist,’ is also problematic
for Millianism. If we suppose that some such statement is true, then ‘X’ fails to denote anything,
but then ‘X’ is meaningless, and a statement which contains a meaningless symbol must itself be
meaningless. Thus it seems no negative existential statement could be meaningful, let alone true,
but surely this is not the case.

Such considerations prompted Gottlob Frege and Bertrand Russell to develop descriptive
theories of reference which in some ways hearken back to Plato’s view. Frege argues that
singular terms can be associated with both “sense” and reference. This richer view of the
meaning of singular terms provides the resources for a solution to Frege’s Puzzle. Problems
raised by negative existential statements are treated differently by Frege and Russell, but both
philosophers analyze these statements in such a way that existence is not a property denied of
some object. These issues are covered more thoroughly in the next chapter, but it is worth noting
that early descriptivist theories were motivated by the most pressing problems afflicting
Millianism.
The basic tenets of descriptivism remained largely unchallenged until Saul Kripke, Hilary Putnam, and David Kaplan, among others, began criticizing the view in the second half of the twentieth century. Descriptivist accounts of reference can have difficulty explaining the functioning of names in intentional and modal contexts. Modal considerations, especially, cast doubt on the possibility of explaining the reference of names and natural kind terms descriptively. Arguments for an alternative account, on which names and natural kind terms refer directly, without any descriptive mediation, have proved convincing.

The problem with this “Kripkean” or “direct-reference” account, at least for some, is that it leads to some unlikely, even unintuitive, conclusions. Numerous long-standing claims and philosophical programs are at least in tension with a Kripkean account of reference: materialism, epistemic internalism, the restriction of possibility to epistemic possibility, and the belief that the function of empirical evidence in determining the truth-value of propositions is to rule out states of the world incompatible with that evidence, to name a few.

According to Scott Soames, a number of philosophers were motivated to undermine the direct reference account in order to preserve their commitments outside the philosophy of language. The first step in their attempt to revive descriptivism was finding descriptions capable of fixing the reference of names and natural kind terms which do not run afoul of Kripke’s semantic arguments. Some manner of rigidifying these descriptions then had to be found, in light of Kripke’s modal argument. Finally, the broad outlines of a “two-dimensional” semantic theory were constructed in order to evade Kripke’s epistemological argument and provide non-threatening accounts of the necessary a posteriori and contingent a priori.

The fundamental two-dimensionalist insight is that there are two distinct ways in which the nature of the world can determine the reference of a term. Reference depends upon both the
context of usage and the circumstances with respect to which a usage is evaluated. Frank Jackson and David Chalmers have developed two-dimensional semantic theories which treat both contexts and circumstances of evaluation as variable, whereas Kripke takes the actual world to be fixed as the context of our usage.

A two-dimensional treatment of contexts and circumstances of evaluation leads into an association of each sentence with distinct primary and secondary intensions. Primary intensions are related to the assignment of propositions in various contexts while secondary intensions are the propositions which are evaluated relative to various circumstances. On some two-dimensionalist accounts, primary intensions serve as arguments of intentional operators, while secondary intensions serve as arguments of modal operators. Because different entities bear modal and epistemic properties respectively, no single proposition can be both necessary and knowable only a posteriori or contingent and knowable a priori.

I think the two-dimensionalist project is misguided. It proves impossible to systematically associate two distinct propositions with a single sentence, one of which is to be the object of intentional attitudes and another which is to bear truth values. Any attempt to do so leads to irresolvable confusion when intentional and modal operators interact in a single proposition. Moreover, there are serious problems which arise when trying to specify the relations of primary intensions to contexts and secondary intensions.

The attempt to revive descriptivism fails even before two-dimensionalism runs into intractable problems. Rigidifying descriptions with the ‘dthat’ operator seems to work, though it is a rather awkward solution. Finding descriptions suitable for rigidification is much more daunting. It appears that all plausible candidates contain either indexical terms or proper names,
the analysis of which leads to confusion. Descriptivism cannot stand, even on the shoulders of Kripke and Kaplan.

Philosophers must abandon descriptivism. The contortions which the ideas of Frege and Russell have been made to undergo in the wake of Kripke’s attacks make the view, in its defensible form, byzantine and unappealing. Though there are certainly pressing problems for direct reference theories, no one would endorse current descriptivist alternatives on their own merits.

Philosophers have been motivated to endorse such views only because of commitments threatened by Kripke’s arguments. It seems to me that discomfort with the necessary a posteriori and fears for the viability of materialism have been the primary motivations for rejecting Kripke’s account of reference. It is admittedly rather baffling at first glance how a proposition which is true in all possible worlds can only be known upon examination of the actual world. Without wading too deep into the arguments concerning materialism, Kripke’s notion of rigidity implies that assertions of identity between objects denoted by rigid designators are, if true, necessary. The necessity of identity lends tremendous force to arguments for dualism from the conceivability of a distinction between mind and body.

As a matter of philosophical methodology, it is these metaphysical and epistemological convictions which ought to be abandoned, not Kripke’s account of reference. It seems to me that the widespread fetish for materialism is simply an overreaction to philosophy’s painfully long history of discussing ethereal fictions. Philosophers feel the painful sting of so many centuries of folly and seek to restore the honor of their tradition, especially in the condescending eyes of scientists, who have so often done exceedingly well for themselves by eschewing discussion of that which is not physical. Hence the clamorous support for materialism.
I think discomfort with the necessary a posteriori has similar psychological origins. The sort of metaphysical necessity involved has a certain air of the mystical which does not sit well with many. But one’s sentiment that things which give off a mystical air are to be avoided is, while admittedly useful at times, not a very sure grounding for argument. Kripke’s claims regarding the reference of names and natural kind terms are exceedingly appealing on their own account and provide solutions to many long-standing problems in the philosophy of language, and so one must follow where they lead, even if it conflicts with widely held sentiments. In fact, where opposing arguments seem equally compelling, it ought to be consciously presumed that the more psychologically inconvenient conclusion is true, given that one’s biases and sentiments are no doubt playing for the opposing team. Such a presumption ought to shore up arguments for the direct reference of proper names and natural kind terms, though they hardly need this bit of support.

Far too much time and energy has been wasted defending descriptivism. This is evident from the baroque apparatus of two-dimensional semantic theory, whose sole purpose is keeping descriptivism alive. Let us accept that names and natural kind terms refer directly to their bearers, and let us unfold the implications of this truth throughout philosophy.
Chapter 2:
Classical Descriptivism

In considering how names are used, it seems clear that they have some meaning, that is, there is some property or conjunction of properties by virtue of which each name is able to serve a communicative function. The most obvious function performed by a name is that of picking out its bearer. This might at first appear to be its only function. There are, however, certain problems which arise when one assumes that the meaning of a name is its referent and nothing more, such as the difficulty in understanding how a name which lacks a genuine referent can still serve a communicative purpose and Frege’s Puzzle. Descriptive theories of meaning were developed by Frege and Russell in response to these problems.

If the meaning of a name is limited to its reference, co-referential names are synonymous and one can be substituted for another in a sentence without any change in the meaning of the sentence. But this seems not to be the case. For example, take the sentence ‘Hesperus is Phosphorus,’ in which ‘Hesperus’ and ‘Phosphorus’ both refer to the planet Venus. It seems that something over and above a statement of Venus’s self-identity is conveyed by this sentence.

Given that the Romans took Hesperus and Phosphorus to be distinct celestial bodies, were I to have told Caesar that Hesperus is Phosphorus, I would have conveyed new information to him. This is inexplicable if it is assumed that ‘Venus,’ ‘Hesperus,’ and ‘Phosphorus’ each have exactly the same meaning, for there would be no difference in the meaning of the tautology ‘Hesperus is Hesperus’ and the apparently informative sentence ‘Hesperus is Phosphorus.’ Caesar would not have a quibbled with the former, but he would have laughed off the latter. These sentences differ, but their names are identical with respect to reference. They must then differ in some other respect.
Frege claims they differ in “sense,” which is “wherein the mode of presentation is contained” (152).\(^1\) It is by means of the sense of a name that a speaker is able to determine its reference. One might say that a sign points to a sense which in turn points to a reference. Frege describes this tripartite relation thus: “The regular connection between a sign, its sense, and its reference is of such a kind that to the sign there corresponds a definite sense and to that in turn a definite reference, while to a given reference (an object) there does not belong only a single sign” (153). A sense directs a speaker to a referent but from a referent one cannot get back to a specific sense. Frege says that a sense “serves to illuminate only a single aspect of the reference” (153). This helps to flesh out what Frege means when he talks about sense as containing the mode of presentation of the referent. A sense seems, for Frege, to be a particular manner of picking out or presenting an object, or, which is the same, a particular manner of conceiving of that object.

Returning to ‘Hesperus is Phosphorus,’ Frege would explain this statement’s cognitive value as stemming from its ability to convey that a single referent (Venus) can be conceived in two distinct ways, thus augmenting the knowledge of someone who had previously conceived of it in only one of these ways. To understand one sense which succeeds in picking out an object is to understand one aspect of that object, and learning another sense by which it can be picked out increases one’s knowledge of the object. Frege says, “comprehensive knowledge of the reference would require us to be able to say immediately whether any given sense attaches to it. To such knowledge we never attain” (153). Senses can be understood as descriptions of particular aspects of their referent. Whether there are aspects of an object which can be known but not captured in

\(^{1}\) All quotations come from Beaney. Where he has left *Bedeutung* untranslated, I translate it “reference.”
a sense is an open question. If there are, it seems these would have to be a sort of “properties” that cannot be captured by language.

Frege’s concept of sense allows for an analysis of identity statements that does not reduce them to mere assertions of a relation between signs. It is tempting to say that the sentence ‘Hesperus is Phosphorus’ asserts only that the name ‘Hesperus’ and the name ‘Phosphorus’ pick out the same thing. Certainly, this is true, and it is a part of what is asserted by the sentence, but such an explanation ignores a crucial aspect of the sentence’s meaning. Understanding the sentence involves more than merely acknowledging the co-referentiality of the signs involved. It requires one to acknowledge that that “star” up there in the evening and that “star” over there in the morning are in fact the same thing. This knowledge is about the planet Venus, not about the names ‘Hesperus’ and ‘Phosphorus.’ The cognitive value of ‘Hesperus is Phosphorus’ cannot be located in its reference, for the tautology ‘Venus is Venus’ has the same reference, and it is not to be found in the relation it asserts between signs, for that is only a small piece of the information conveyed. It must be located somewhere in between sign and reference, where Frege finds his sense.

In this space, between sign and reference, is where Frege’s locates thought, understood not as “the subjective performance of thinking but its objective content” (156). Frege argues that because the substitution of co-referential terms in a sentence changes the thought expressed by the sentence, the thought cannot be associated with the reference of the sentence. For example, the sentences ‘Mark Twain was Samuel Clemens’ and ‘Mark Twain was Mark Twain’ clearly express different thoughts and yet they have the same reference. If there is any doubt as to whether or not these sentences express different thoughts, one can easily imagine someone acknowledging that Mark Twain was himself while doubting that he was Samuel Clemens. It is
hard to see how this could be so if one and the same thought were expressed by both sentences. Frege concludes that “the thought, accordingly, cannot be the reference of the sentence, but must rather be considered as its sense” (156).

If the thought expressed by a sentence is its sense, then a question arises as to the nature of its reference. A sentence composed of genuinely referential terms would certainly seem to have some reference itself. Frege discusses literature as a paradigmatic case of language which has sense but lacks reference. Literature is no different from more mundane forms of communication in its ability to convey thought. It differs in that it is ridiculous to ask of a piece of literature whether it is true or false, whereas we make such judgments constantly with regard to other uses of language. Frege concludes, “It is the striving for truth that drives us always to advance from the sense to the reference” (157).

Frege establishes the link between reference and truth value. One cannot be had without the other. But he goes further:

We have seen that the reference of a sentence may always be sought, whenever the reference of its components is involved; and that this is the case when and only when we are inquiring after the truth value.

We are therefore driven into accepting the truth-value of a sentence as constituting its reference. By the truth-value of a sentence I understand the circumstance that it is true or false (157).

Taking the reference of a sentence to be its truth value seems unobjectionable in so far as no problems immediately arise from this result and no alternative suggests itself. Yet I must admit I find the suggestion that a sentence refers to either the True or the False very difficult to understand.
Frege says that “the sense of a proper name is grasped by everybody who is sufficiently familiar with the language” (153). Frege allows for some variation in the senses associated with a particular name by different speakers so long as each picks out the same referent, but this is still a striking claim. Frege seems to require every speaker who uses a name to be able to come up with at least one uniquely referring description that picks out the referent of the name. This flies in the face of common experience, for it certainly seems as though people use names all the time for things which they would be unable to pick out by description. For example, my inability to describe a gasket does not seem to impede my use of the word. Without knowledge of senses that pick out a name’s referent, Frege has no explanation of how a speaker could successfully use a name, yet this seems to happen all the time.

This theory elegantly solves Frege’s Puzzle. It is also provides an explanation of negative existential statements and, more broadly, the usefulness of terms which fail to refer, though not in so elegant a fashion. Frege acknowledges that such terms occur: “It may perhaps be granted that every grammatically well-formed expression figuring as a proper name always has a sense. But this is not to say that to the sense there also corresponds a reference” (153). He also says that “it is of the reference of the name that the predicate is either affirmed or denied” (157). He thus seems compelled to say that any statement which contains a non-referential term lacks a truth value, or at least that its truth value cannot be known, for if one can neither affirm nor deny the predicate of the subject, presumably one cannot judge the truth of the sentence.

Frege solves this apparent problem by, in the case of negative existential statements, analyzing existence as a second-level concept. The statement ‘The Easter Bunny does not exist’ is analyzed ‘—(∃x) (x = is the Easter Bunny),’ which is to say that the first-level concept expressed by ‘is the Easter Bunny’ does not fall under the second-level concept expressed by the
existential quantifier. There is no longer any subject of which to affirm or deny the property of existence, which has itself dropped out in the analysis, so the problem which seemed to arise from 'the Easter Bunny' lacking reference is avoided.

Negative existential statements are more problematic for Millian accounts of reference. The problem is that a name denoting something which does not exist fails to refer, but if there is no more to its meaning than its reference, such a name will be meaningless, and it seems that a sentence containing a meaningless name will itself be meaningless. But meaningful negative existential statements abound.

Russell, like Frege, analyzes denoting phrases into a function component and an argument component: “I use ‘C (x)’ to mean a proposition 2 in which x is a constituent, where x, the variable, is essentially and wholly undetermined” (1905, 480). In a footnote, he explains that by “proposition” he means “more exactly, a propositional function” (1905, 480). He goes on to list possible categories (corresponding to Aristotle’s distinctions among affirmations of predicates of all, none, or part of a subject), which these propositional functions may fall into: C (x) is always true, ‘C (x) is false’ is always true, and ‘it is false that ‘C (x) is false’ is always true.’

Russell thinks that with these categories, grounded in the primitive notions ‘false’ and ‘always true,’ and his functional analysis he can get rid of denoting phrases and so avoid the problems to which they give rise. He argues that “denoting phrases never have any meaning in themselves, but that every proposition in whose verbal expression they occur has a meaning” (1905, 480). For example, ‘all men are mortal’ is analyzed ‘‘for all x, if x is a man, then x is mortal’ is always true’ When some object is assigned to x, the propositional function will determine a truth value, and the denoting phrase ‘all men’ drops out in the analysis.
Russell makes a distinction between primary and secondary occurrences of referential terms. In an independent clause, a term has “primary” occurrence, while in a subordinate clause it has “secondary” occurrence. The analysis of a proposition in which a referential term (or “denoting phrase,” in Russell’s terminology) has primary occurrence will differ from one in which it has secondary occurrence. Often the same sentence can be interpreted in different ways, depending on whether a term is taken to have primary or secondary occurrence. For instance, Russell shows that the statement ‘the present King of France does not exist’ can be analyzed in two distinct ways. If ‘the present King of France’ is taken as primary, it is analyzed ‘there is some $x$ such that $x$ is the present King of France, and $x$ does not exist.’ On this analysis, the statement contradicts itself, for if there is some $x$, then that $x$ exists. The alternative analysis, which takes ‘the present King of France’ as secondary, is ‘it is not the case that there is some $x$ which is the present King of France and does not exist.’ This is true because there is no $x$ which is the present King of France, and, happily, this seems what we would expect a speaker to have intended to convey by an utterance of this sentence. Russell’s analysis of propositions in conjunction with his distinction between primary and secondary occurrence allow for a satisfactory analysis of negative existential statements.

Russell solves Frege’s Puzzle by arguing that names are abbreviated descriptions: “Common words, even proper names, are usually really descriptions. That is to say, the thought in the mind of a person using a proper name correctly can generally only be expressed explicitly if we replace the proper name by a description” (1917, 156). He would explain the cognitive value of ‘Hesperus is Phosphorus’ as stemming from its claim that distinct descriptions pick out the same referent. This is quite similar to Frege’s explanation in terms of sense. The similarity continues in that Russell assumes different speakers may associate different descriptions with the
same name, and that this is inconsequential so long as they have the same referent. However, unlike Frege, who leaves open the possibility that a name might have reference but lack descriptive sense, Russell is committed to analyzing all names (with the exception of the “logically proper” names ‘this,’ ‘that,’ and ‘I’) as descriptions.

There is a fair bit of common ground between Frege and Russell, enough that one is justified in speaking of them together as descriptivists. They are in agreement that most proper name can be replaced by some suitable description without any change in the meaning of the sentence in which it occurs, that it is descriptive knowledge which facilitates reference, and that the meaning of a name is rarely limited to its reference. Most of Russell’s ideas, including his distinction between primary and secondary occurrence, are compatible with Frege’s theory. It is thus possible to construct a composite description theory that combines the best of each approach. Classical descriptivism should be understood as such a composite, one which relies heavily on Frege’s conception of sense, thought, and indirect discourse, as well as his discussion of subjective ideas and the reference of sentences but includes Russell’s method of analysis and the notions of primary and secondary occurrence.
Classical description theories drawing on the work of Frege and Russell are able to provide explanations of most linguistic phenomena and do a convincing job dealing with the problems which thwarted purely referential theories. For decades, there was little impetus to question the central tenets of descriptivism until Kripke delivered a powerful series of counterarguments in 1970, later compiled in *Naming and Necessity*. Most notably, he pressured descriptivist accounts of meaning by examining sentences occurring inside various modal and intentional operators. His arguments possess tremendous intuitive force and put any description-based theory of meaning on the defensive.

Kripke argues against the strong descriptivist claim that the meaning of a proper name or natural kind term can be identified with a co-referential description or cluster of such descriptions. If a name and its associated description(s) were synonymous they could not come apart in any possible world. If some object fulfilled the description(s), that object would necessarily be the referent of the name, and if the term referred to some object, that object would fulfill the associated description(s). This is clearly false.

One might associate the name ‘Abraham Lincoln’ with descriptions such as ‘the man who delivered the Gettysburg Address,’ ‘the first Republican President,’ or ‘the man who debated Stephen Douglas.’ For the strong descriptivist, such descriptions are synonymous with ‘Abraham Lincoln.’ But Lincoln might never have run for office and so never have been in a position to debate Douglas or deliver the Gettysburg Address. Moreover, with Lincoln absent from the political scene, some other Republican would presumably have been the first of his party elected President, but that man would not have been Lincoln. Intuitively, it seems our use
of ‘Abraham Lincoln’ refers to a particular man even in situations in which that man does not satisfy any of the descriptive conditions we commonly associate with him, given the actual course of history. If this intuition is correct, then, contra strong descriptivism, the name is not synonymous with any immediately plausible co-referential description(s).

In the context of these modal arguments, Kripke presents his notion of rigidity. Kripke calls something a rigid designator “if in every possible world it designates the same object,” provided the object exists (48). He calls a designator “strongly rigid” if its referent exists in all possible worlds. Soames suggests an intuitive test of rigidity: “a singular term t is a rigid designator iff the individual who is t could not have existed without being t, and no one else who is not the individual who is t could have been t” (16).

Kripke claims that proper names and natural kind terms are rigid designators. Given their rigidity, one can easily see why true identity statements involving such terms are necessary. Suppose ‘Cicero’ rigidly designates some man and ‘Tully’ rigidly designates some man. If it is true that Cicero is Tully, then it is true in all possible worlds that Cicero is Tully because the reference of both terms is constant across possible worlds.

At this point, a word should be said about natural kind terms. I have been using the terms ‘name’ and ‘natural kind term’ without particular care. This shouldn’t be problematic, as Kripke’s arguments apply equally well to both sorts of terms, but there are important distinctions between them. Names are singular terms—each name refers to a single object. Natural kind terms refer to classes of objects, specifically classes comprised of all objects of a particular kind found in nature. To see how this works, let’s look at Kripke’s account of the kind ‘cat’: “The original concept of cat is: that kind of thing, where the kind can be identified by paradigmatic instances. It is not something picked out by any qualitative dictionary definition” (122). Kripke
contends that kind terms, like names, are directly referential. Natural kind terms give rise to many interesting instances of the necessary a posteriori (‘water is H₂O’) because we are in a better position to discover interesting a posteriori truths about natural kinds than the bearers of proper names. For the purposes this paper, the differences between these sorts of terms is of minimal relevance.

It is not only certain corollaries of the description theory concerning modal truth values whose falsehood is demonstrated by the Lincoln example. The strong version of the description theory implies that, since the meaning of a proper name is a suitable conjunction of co-referential descriptions (to be abbreviated D∗), whenever anyone takes an intentional attitude—belief, for example—toward a proposition which predicates something of an object designated by a proper name (x believes that n is F), he must also hold the same attitude toward the proposition which predicates that property of the same object designated by an appropriate description (x believes that D∗ is F). The two sentences, ‘n is F’ and ‘D∗ is F’ express the same proposition. One cannot believe one and not the other, for there is only a single proposition in question. But clearly it is possible for someone to simultaneously believe that the first Republican President was born in a slave state while harboring no such belief about Abraham Lincoln.

The description theory’s epistemological problems are not limited to intentional attitude ascriptions but also concern supposedly synonymous sentences which differ in epistemic value. If the name ‘Abraham Lincoln’ is synonymous with ‘the man who delivered the Gettysburg Address’ then I can know a priori that Lincoln delivered that address. But I can’t know this without first looking into the matter. My knowledge that Lincoln did in fact deliver the Gettysburg Address is founded upon consultation of books written by historians who learned this fact from other historians who ultimately learned it from historians who had examined evidence,
a contemporary newspaper perhaps, indicating that it was Lincoln who gave the speech. The
episocm chain could be traced back even further to the men who heard the speech, but the point
is clear enough that my knowledge that Lincoln gave this speech is grounded in empirical
evidence. I might conceivably be wrong in believing that Lincoln gave this speech. Perhaps he
pawned it off on a subordinate and, after it received such a tremendously positive response, paid
off the journalists in attendance to tweak the story, so that he might be known to posterity as the
man who delivered the Gettysburg Address. Knowledge that Lincoln delivered this speech is a
posteriori, though the description theory implies its a priority. These arguments from intentional
attitude ascriptions and epistemic properties apply only to the strong version of the description
theory which claims that names are synonymous with descriptions.

Descriptivism implies a two-way necessary relation between the reference of a name and
its descriptive senses. If a description or cluster of descriptions (given the variety of descriptive
senses which different speakers may at different times associate with a particular name)
determines the reference of a name, that name would fail to refer if the description(s), or at least
a sufficient number of them, were not fulfilled. Conversely, whatever object fulfills the
appropriate description(s) is the referent of the name.

Kripke points out, in what is known as his semantic argument, that, contrary to the
requirements of the description theory, names refer even in counterfactual scenarios in which
their supposedly reference-fixing descriptions are not fulfilled, provided their referents exist. For
example, let’s assume that ‘the man who gave the Gettysburg Address’ fixes the reference of
‘Abraham Lincoln.’ Returning to the hypothetical in which Lincoln does not actually deliver that
address but only takes credit after the fact, what do our intuitions tell us about the reference of
‘Abraham Lincoln’ in that scenario? It seems our use of the name still refers to the familiar
bearded emancipator, but on the fix-the-referent version of descriptivism, we actually refer to his eloquent subordinate. This gets things wrong. It seems all the descriptive facts which I believe to be true of Lincoln might be false without any alteration in the reference of the name ‘Abraham Lincoln.’

It should be noted that none of the above claims are threatened by the obvious truth that some other man might have been named ‘Abraham Lincoln.’ In such a world, the inhabitants would no doubt use the name to refer to that other man, but our actual use of the name ‘Abraham Lincoln’ would not refer to that other man, but necessarily to the familiar bearer of that name in the actual world, regardless of what he might have been called in any counterfactual scenario under discussion.

If this claim seems doubtful due to the possibility of interruptions of the relation of the man we call ‘Abraham Lincoln’ to his name, it is only due to confusion in the interpretation of the counterfactual scenarios under consideration. Kripke’s arguments present for speculation various counterfactual accounts of a particular man’s life and then ask if his name, when used in the actual world, refers to the man in these accounts. The foundation of contemporary two-dimensional semantic theories is laid in the confusions which arise from the consideration of various states of affairs as, respectively, actual and counterfactual.
Chapter 4: 
Finding the Right Descriptions

Kripke’s semantic argument convincingly demonstrates that the reference of most names does not depend on their fulfilling descriptions with which we commonly associate them. Kripke also presents related problems by exposing how frequently the reference of a name differs from that of its commonly associated co-referential descriptions in modal contexts. Given these arguments, any attempt to revive descriptivism must employ descriptions whose fulfillment is necessary for the use of the name and whose reference does not vary across modal contexts.

Any description which is to fix the reference of a name is one which must be satisfied in order for use of the name to be possible, for if a reference-fixing description failed to refer so would its associated name. The typical approach for constructing such descriptions is to use Kripke’s own idea of a causal-historical chain of reference transmission.

Kripke suggests that names typically acquire reference through a process beginning with some sort of “initial baptism” in which an authority determines that some object shall be referred to by such and such a name, new parents naming their baby, for example. Soon other speakers learn the name and its intended reference from these authorities. They use the name with the intent that its reference shall accord with the baptizers’ usage. Other speakers eventually learn the name from them, and so on, all speakers intending their reference to accord with the community’s. It is not necessary that anyone know the details of the initial baptism or from whom they learned the name because of the shared assumption that everyone’s usage is intended to accord with everyone else’s.

If a name refers by virtue of some causal connection it bears to both a speaker and its referent, then, descriptivists have noted, it should be possible to describe these causal
relationships. I might fix the reference of ‘Obama’ with a description such as ‘the person I have heard of under the name ‘Obama.’’ Here I effectively borrow the reference of the name from some other speaker, who is himself presumably borrowing reference from some third person, creating a chain which, one would hope, reaches back to the man of which we wish to speak.

This account is appealing because the causal-historical theory of reference transmission is well founded and at first glance there seems no reason why the relations constituting the causal-historical chain cannot be described. But there are several problems with using descriptions such as the one above to fix reference. There is no certainty that descriptions like ‘the place of which I have heard under the name ‘Manzanita’’ will lead me back to the intended referent. I may have heard the name ‘Manzanita’ from someone who had confused it with ‘Mazatlan,’ and thus, according to the theory, wind up referring to that coastal Mexican city every time I use the name ‘Manzanita.’ But of course it seems I would still refer to Manzanita even if I had heard the name from someone who was confused about its reference.

Misuse is not the only problem. The intended reference of a name can change. If we suppose that I fix the reference of Madagascar by the description ‘the causal source of this token ‘Madagascar,’’ we must conclude that the name refers to the capital of Somalia. The name did not originally refer to the island now known as Madagascar. It was a corruption of the name ‘Mogadishu,’ which Marco Polo mistakenly took to refer to the island now known as Madagascar. Ultimately, the causal source of the name is an East African port city to which no one currently using the name ‘Madagascar’ intends to refer. Changes of this sort are not uncommon, and any descriptive attempt to fix reference by finding the end of a causal chain of usage is unable to account for historical changes in the intended reference of a name.
The previous example illuminates one respect in which reference depends on social context. If I had been in East Africa shortly after Marco Polo’s travels and used the name ‘Madagascar’ while speaking with the natives, they would naturally have assumed I was intending to refer to Mogadishu, and thus I would in fact have referred to Mogadishu, unbeknownst to me. Back in Venice, of course, I would have referred to the island rather than the port because that is what my audience would have taken to be the intended reference of my usage of ‘Madagascar.’ It is not only the intentions and knowledge of the speaker, but also those of his audience, upon which reference depends.

Another example emphasizing the importance of a speaker’s audience is given by Jonathan McKeown-Green: If one knows that in every town in a particular region of Ireland there is always exactly one man named Patrick O’Grady, one could walk into the pub in any of these towns and ask for Patrick O’Grady, each time referring successfully to some particular local, without ever having heard anything about him. Here it is highlighted how little anyone actually relies on his own introduction to a particular name. The reference of a particular usage of ‘Patrick O’Grady’ depends on assumptions held by the audience, as well as the assumption that the speaker intends his usage to accord with theirs.

As Soames points out, names already have meaning before I encounter them, and when I incorporate them into my idiolect, I do so with the intention of retaining this public meaning. Causal-historical descriptivism adheres to the letter of Kripke’s ideas but fails to incorporate the social spirit of his thought. The problem which leads to all these difficulties lies in associating a name with a personalized descriptive sense which is only connected to others’ usage at the periphery. Language users are not nearly so isolated. The meaning of a name is not only social in the sense that each individual can ultimately find his way back to the same reference. It is rooted
in an expectation that each speaker intends to use names in accordance with their public meaning.

The above arguments throw into serious doubt the possibility of finding descriptions for most names which unfailingly refer to their intended objects in the actual world. There are certainly some circumstances in which it is possible to fix the reference of a name descriptively. To use an example of Kripke’s, the name ‘Neptune’ was introduced expressly for the purpose of designating whichever object was causing specific perturbations in the orbit of Uranus. The name ‘Neptune’ referred by stipulation to whatever fulfilled this description. But it is still not the case that the reference of such a name can be identified with the description because it is perfectly intelligible to suppose, for instance, that had some other object been altering the motion of Uranus, Neptune would never have been called ‘Neptune.’ We might have called some other body ‘Neptune,’ but it would not be Neptune because names are rigid designators. Even in those cases in which reference is fixed descriptively, it is fixed by some contingent feature of the actual world whose obtaining in any counterfactual circumstance in which the name is used is irrelevant to its reference. The only way around this argument is to somehow rigidify the reference-fixing description(s) so that it acts like the name whose reference it fixes in counterfactual situations.

There are certainly descriptions which function in this way, but they are not the sort of descriptions which one would likely hear from passers-by upon asking them for the meaning of some name. Most of descriptivism’s appeal is lost in this process. It is no longer the case that one can use a name by virtue of knowing some description(s) with which the bearer is commonly associated but only by virtue of implicitly knowing some obscure rigidified description.
There are two means of constructing rigidly designating descriptions which might conceivably fix the reference of a name: the actuality-operator and the dthat-operator. Each of these approaches rigidifies by anchoring a description to the actual world in a way which makes it immune from the usual effects of modal operators. The actuality-operator does this in a straightforward fashion, simply stipulating that any description it governs is always to be evaluated with respect to the actual world. The dthat-operator does so demonstratively. It is analogous to ‘that,’ but by definition designates rigidly. Once I refer demonstratively to some object using ‘dthat,’ the associated description continues to pick out that same object in any possible world.

The first approach utilizing the actuality-operator, for all the appeal of its straightforwardness, is unworkable. Soames points out that if the reference of names were fixed by actually-rigidified descriptions, thinkers in counterfactual circumstances, i.e. non-actual possible worlds, would need to have beliefs about the actual world in order to use names. Suppose some thinker, Erika, in some non-actual possible world believes the proposition expressed by ‘Cicero was bald.’ Because, on this account, the reference of ‘Cicero’ is fixed by a description such as ‘the x: actually Cx,’ the actual world is a constituent of the proposition which Erika believes. So Erika, in a world in which Cicero never lost his hair, who believes the proposition expressed by ‘Cicero was bald,’ believes something true about a man in another possible world (the actual world).

One would imagine that thinkers in other possible worlds are perfectly capable of holding beliefs about those worlds, and it is not at all clear why any beliefs containing names which they hold must be about the actual world. In fact, it is highly doubtful that a speaker in some non-actual possible world could have beliefs about the actual world. She could presumably
have beliefs about what is true with respect to the set of possible worlds and thus beliefs about
the actual world *qua* possible world, but she could not have beliefs about the actual world *qua*
actual world, for how would she know which world is actual? For the actuality-operator to serve
its purpose, anyone using an actualized description must not only believe the relevant description
with respect to the actual world but must hold a belief about the actual world *qua* actual world.
Of course, the descriptivist argument requires more than the possibility that agents in non-actual
possible worlds could have beliefs about the actual world. It requires that such beliefs are a
necessary condition for their use of names, which is absurd.

It seems possible to rigidify descriptions using the dthat-operator in a way which averts
unwanted consequences of Kripke’s modal and semantic arguments, but in doing so one must
accept some troubling results. Agents in non-actual possible worlds need not believe anything
about the actual world in order to hold beliefs containing dthat-rigidified descriptions because
they are perfectly capable of designating some object ‘x’ demonstratively, just as someone in the
actual world is able to designate that same ‘x’ demonstratively, making it possible to refer to the
same object across possible worlds without any epistemic contact between worlds. The dthat-
operator thus avoids the pitfall exposed by use of the actuality-operator, but there are other
difficulties. Because any description governed by the dthat-operator is, as demonstrative, directly
referential, a descriptivist theory which makes use of the operator faces all of the problems which
confront any Millian view, including Frege’s Puzzle and the problem of negative existential
statements. One wonders whether this is too high a price to pay in order to salvage the claim that
the reference of proper names is fixed descriptively, for the appeal of this now complicated
theory is no longer clear.
Kripke’s semantic argument precludes the most intuitively plausible descriptions from fixing the reference of names. Acceptance of a causal-historical theory of reference transmission compels any reference-fixing description(s) to somehow refer back to the original introduction of the name into the language. It is extremely difficult if not impossible to construct such descriptions, but even if this possibility is granted, Kripke’s modal argument poses further problems. Because names are rigid designators, any description which purports to fix the reference of a name must rigidly designate its bearer. A description can be made to do this by means of a suitable actualizing operator, but due to problems involved in the use of such operators, the only possible approach requires turning reference-fixing descriptions into directly referential demonstrative expressions. In the pirouette necessary to avoid Kripke’s semantic and modal arguments, much of descriptivism’s original appeal is lost, and while possible in some cases, it is doubtful that viable reference-fixing descriptions can be constructed for most names.
Chapter 5:  
Historical Roots of Two-Dimensional Semantics

I will temporarily put aside the discussion of descriptivism in order to sketch a two-dimensional view of semantics on which contemporary descriptivist arguments lean. This chapter will trace various lines of thought which, when extended, arguably beyond what can be justified, comprise semantic two-dimensionalism.

David Kaplan, in “Demonstratives,” unobjectionably analyses linguistic meaning as comprised of two distinct components: content and character. For sentences, this distinction is fairly straightforward: “The content of a sentence in a context is what has traditionally been called a proposition” (501). Frege, as Kaplan notes, would have called the content of a sentence a thought, which is the same as the sense of the sentence. The content of a sub-sentential expression is what that expression contributes to the proposition.

The distinction between contexts of utterance and circumstances of evaluation is critical for Kaplan. A context of utterance is fairly easy to understand. Knowledge of a context must include, at least, a speaker, a time, a place of utterance, and, for demonstratives, knowledge of the associated demonstration. Circumstances of evaluation are “actual or counterfactual situations with respect to which it is appropriate to ask for the extensions of a given well-formed expression” (502).

Character is a function from contexts of utterance to contents. Kaplan’s analysis is especially clear in the case of indexical terms. The character of the term ‘I’ in the sentence ‘I am here now’ is a function which fixes whoever utters the sentence as the content of the term ‘I.’ The picture here is that the character of, say, a sentence is a function which maps the sentence
from a context of utterance to a content (proposition) which is then evaluated with respect to some particular circumstance in order to determine a truth value.

Kaplan’s analysis of meaning into the two components of content and character is sometimes referred to as two-dimensional. This sort of two-dimensionalism is unproblematic. Soames refers to it as “benign” in contrast to the “ambitious” two-dimensionalism of authors like Jackson and Chalmers. Where I speak of “two-dimensionalism” I intend to refer only to the more ambitious varieties.

Kaplan analyzes rigid designators in this framework: “Rigid designators (in the sense of Kripke) are terms with a Stable Content” (10, XIX). It does not matter with respect to which circumstances a rigid designator is evaluated, as it will always designate the same object. That is why it is said to have “Stable Content.” In fact, Kaplan says, “All directly referential expressions (as well as all rigid designators) have a fixed content. [What I elsewhere call a stable content.]” (502).

By contrast, a term has “Stable Character” if its character determines the same content in all contexts of utterance. Most expressions have Stable Character. All proper names, for instance, have Stable Characters as well as Stable Contents. Indexicals and demonstratives are notable exceptions. They have characters which determine different contents depending on the context of utterance.

So far there is nothing with which to take issue in Kaplan’s account, but he goes further in a way which sets up the two-dimensionalist position. Kaplan takes logical truth to be a form of a priority (538). He then rightly points out that “a truth of the logic of demonstratives like ‘I am here now’ need not be necessary” (538). Kaplan explains how a statement can be both logically true and contingent by suggesting, in a way which foreshadows the two-dimensionalist
distinction between primary and secondary intensions, that modal and logical properties inhere in different entities:

\[E. \text{ Corollary 3} \] The bearers of logical truth and of contingency are different entities. It is the character (or, the sentence, if you prefer) that is logically true, producing a true content in every context. But it is the content (the proposition, if you will) that is contingent or necessary (539).

This is especially troubling given his claim that logical truth is a species of the a priori. Kaplan has laid the groundwork for a semantic analysis which bisects the meaning of sentences into distinct entities, one which is relevant in modal contexts and another which is relevant in epistemic contexts. His basic insights concerning content and character are sound, and they are indispensable to any analysis of indexical reference, but corollary 3 and his claims about logical truth and a priority go too far.

Robert Stalnaker attempts to explain assertion by constructing a two-dimensional model of conversation. He asserts that the purpose of conversation is to locate the actual world among the set of possible worlds. Conversational participants first agree (this sort of agreement is typically implicit) upon a set of beliefs, which limits the set of possible worlds. Only those worlds consistent with the shared beliefs of the interlocutors are candidates for actuality. Stalnaker refers to this limited set of possible worlds as the “context set.” As conversation continues, the goal is to locate the actual world among the worlds of the context set. To this end, assertions are made which propose additional beliefs. If all participants in the conversation acquiesce to an assertion, the context set is narrowed, all worlds which are incompatible with the augmented set of shared beliefs being removed from the set.
This model forces Stalnaker to place various restrictions on the sort of propositions which can legitimately be asserted. He writes, “1. A proposition asserted is always true in some but not all of the possible worlds of the context set” (89). The rationale for this restriction is easy to see. Any assertion which is false with respect to all worlds of the context set is one which delegitimizes the entire conversation. If new beliefs are arrived at which are incompatible with the entire context set, the actual world cannot be among those of the set, and the conversation cannot continue without some revision of shared beliefs and enlargement of the context set so as to bring the new assertion into accord with Stalnaker’s first principle. A proposition cannot be true with respect to all worlds of the context set because that would make it trivial. It would not allow for any narrowing of the context set and so would not advance the conversation.

Stalnaker’s second principle is “2. Any assertive utterance should express a proposition relative to each possible world-state in the context set, and that proposition should have a truth value in each possible world-state in the context set” (89). The rationale for this principle is straightforward. If it were not possible to determine the truth value of a proposition expressed by an assertion with respect to some world in the context set, it would be impossible to determine whether or not to retain that world in the set. Such an assertion would fail to perform its proper function in the conversation, and an incompletely determined context set would render conversation, in Stalnaker’s sense, impossible because it would always remain unclear whether or not the context set had been winnowed down to a single, actual world.

Stalnaker’s third principle—“3. The same proposition should be expressed relative to each possible world-state in the context set”—is the most problematic (89). It is clear why he requires such a principle. The conversational participants do not know which world is actual, which world they are in. If an assertion expressed different propositions in different world-states
taken as contexts of utterance, they would not know which proposition was expressed in their own world-state and so would be unable to determine with which possible states of the world it was compatible. In such a circumstance, the context set would be indeterminate, which is unacceptable.

An assertion which seems to violate any one of these principles can typically be interpreted in such a way that the violation is only apparent. If a speaker asserts something which is false with respect to every world of the context set, his interlocutors might realize they had been assuming too small a set and revise accordingly. Or a speaker might not actually mean what he seemed to assert, as when a literal interpretation of an utterance is inappropriate. In fact, a speaker’s knowledge of the context set may allow him to make ambiguous assertions, assuming that his fellows will interpret him in a way that conforms to the context set.

There are, however, certain commonplace assertions which seem difficult to square with Stalnaker’s principles, even granting this interpretive license. Assertions of necessary truths, for example. If a speaker, pointing at a certain man, says “That is either John Kerry or Joe Biden,” his assertion is, depending on the context, either necessarily true or necessarily false. The assertion, which seems unobjectionable, and even informative, does not accord with Stalnaker’s first principle. To make sense of this, Stalnaker employs the notion of a “propositional concept,” which is “a function from possible worlds into propositions” (81). This can be expressed by a truth table. The propositional concept associated with the above assertion is, where ‘a’ is a world in which the man pointed out is John Kerry, ‘b’ a world in which he is Joe Biden, and ‘c’ a world in which he is neither man
The columns represent possible worlds as circumstances of evaluation, while the rows represent possible worlds as contexts of utterance. It is possible to think of this table as a function from possible worlds into propositions because Stalnaker takes propositions to be “function[s] from possible worlds into truth values,” with the possible worlds here referred to taken as circumstances of evaluation (81).

Because the assertion contains a demonstrative, the proposition it expresses depends on the context in which it is uttered, which is to say, it expresses different propositions in different contexts, in violation of principle 3. Uttered in contexts \(a\) and \(c\), the assertion is trivial (assuming names are rigid designators). Uttered in context \(c\), it is a necessary falsehood. This conflicts with Stalnaker’s first principle. He resolves this dilemma by reference to the “diagonal proposition” associated with the assertion. The diagonal proposition is the proposition that would be true in a context if the proposition were asserted in that context. This can be read along the diagonal of the table from top left to bottom right.

There is some vagueness in the notion of a diagonal proposition. It is not altogether clear exactly which proposition is being spoken of. I think this vagueness stems from philosophical error on Stalnaker’s part. The same error is carried into the two-dimensionalism which makes use of Stalnaker’s notion. The underlying problems will be addressed in the discussion of Soames’ criticisms of two-dimensionalism. The difficulty in pinning down talk of diagonal
propositions foreshadows strong two-dimensionalist ambiguity in the notion of a primary intension.

The diagonal proposition associated with the assertion analyzed in the truth table above is neither true with respect to all worlds of the context set (and thus trivial) nor false with respect to all of them (and thus illegitimate). It seems capable of performing the proper role of assertion, which is narrowing the context set. It also seems to achieve the intuitively correct result of ruling out all worlds in which the man demonstrated in the assertion is neither John Kerry nor Joe Biden.

Stalnaker’s model foreshadows two-dimensionalist attempts to explain apparently a posteriori necessary truths by varying assumptions as to both which world is actual and which world is counterfactual, in effect varying both contexts of utterance and circumstances of evaluation. Stalnaker no doubt takes his diagonal propositions to accord with his third principle, but the question which needs answering here is which single proposition is it exactly which is expressed at all these various worlds of the context set? It certainly seems as though different propositions are expressed in different contexts, and how could it be otherwise given the demonstrative in the assertion?

Stalnaker’s model provides an even less convincing account for assertions of necessary a posteriori truths not containing indexicals or demonstratives. Take an assertion such as ‘Hesperus is Phosphorus.’ If these names are rigid designators, then the proposition expressed by this sentence is the same in every context of utterance. In fact, it is true regardless of which possible world is taken as a context of utterance and which is taken as a circumstance of evaluation. Its propositional concept is
Even the diagonal propositional is a necessary truth. In order to bring such an assertion into accord with his first principle, Stalnaker would have to somehow make the reference of the names ‘Hesperus’ and ‘Phosphorus’ depend upon the context of utterance, which brings us to full-blown ambitious two-dimensionalism.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>b</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>c</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>
Chapter 6: Two Varieties of Two-Dimensionalism

According to ambitious two-dimensionalism, every sentence S is associated with two distinct intensions: a primary intension which is a proposition that is true at every context, and only those contexts, to which the character of S assigns a proposition which is true in that context and a secondary intension which is the proposition assigned by the character of S to a particular context. Beyond this analysis of sentences, there are two distinct approaches to filling out a two-dimensionalist framework, which Soames refers to as, respectively, the “strong” and “weak” variants of the theory. The most salient difference between the two is that whereas strong two-dimensionalism denies that any proposition can be both necessary and knowable only a posteriori or contingent and knowable a priori, weak two-dimensionalism grants these possibilities while explaining them in a deflationary way.

To better understand the distinction between primary and secondary intensions and its relation to descriptivism, it is helpful to look at an example. Let’s assume the reference of the name ‘Aristotle’ is fixed by the description ‘the teacher of Alexander the Great.’ Let’s take a to be a world in which Aristotle taught Alexander the Great and was fond of dogs, b to be a world in which Aristotle taught Alexander the Great and hated dogs, and c to be a world in which Thales, who hated dogs, taught Alexander the Great while Aristotle did not but was fond of dogs. I will present the truth values associated with the primary and secondary intensions of the sentence ‘Aristotle was fond of dogs’ in a table. The columns represent worlds taken as circumstances of evaluation while the rows represent worlds taken as contexts of utterance.
The truth values indicated in the boxes represent the truth value of the secondary intension expressed in a given context and evaluated with respect to a particular world. The primary intension can be identified with the Stalnaker-style diagonal propositional running from the top left to the bottom right.

It is difficult to understand exactly which proposition this is. Essential to strong two-dimensionalism is an analysis of meaning into two components, analogous to Kaplan’s distinction between character and content. But character is a function. It is not the sort of thing which can bear epistemic properties. In developing the notion of a primary intension, the strong two-dimensionalist needs to retain the basic features of character while explaining how primary intensions can be the bearers of a priority and a posteriority.

Primary intensions are thus turned into propositions, and since a proposition must express some thought, we are told that primary intensions are propositions which assert the truth of the proposition that the character of the sentence assigns to the context. This comes to look like Stalnaker’s diagonal proposition because, given that primary intensions always express the truth of the associated secondary intensions only with respect to the world of the context, it only makes sense to evaluate them with respect to the world of the context.

But this is all rather tacked on. At first glance, it seems ludicrous to analyze a sentence as comprised of two distinct propositions one of which simply asserts that the other is true in the
context of utterance. It's not even clear that an assertion of the form ‘p is true’ asserts anything more than ‘p.’ This additional proposition conveys no information over and above p and simply tracks p’s truth value in different contexts. The only motivation for contriving such a notion is that it can play the role of bearing epistemic properties, which is required for strong two-dimensionalism.

While both strong and weak two-dimensionalism claim that all sentences are associated with both primary and secondary intensions, there are important differences in the way these two views characterize primary intensions. Though a strong two-dimensionalist might in principle take primary intensions to be simply Kaplan-style characters, typically they are claimed to be propositions defined in terms of such characters. The strong two-dimensionalist is committed to characterizing primary intensions in such a way that they are invariant across all possible worlds even for sentences, such as those containing indexicals, whose characters are not constant functions. As a proposition, a primary intension must make some assertion, and the only assertion which seems to accord with the requirements outlined above is that of asserting that the character of its associated sentence S expresses a truth.

It would not do, however, to assert that the character of S expresses a truth in a particular context, for primary intensions must remain invariant across contexts. A primary intension must assert that the character of S expresses a truth unrelativized to any context or set of contexts. It is coherent to assert that a character, which is a function, is true without specifying any context as argument so long as this is taken to mean only that the character is true in any context at which it assigns a true proposition and contexts are identified with world-states. Again, it should be noted that the primary intensions of strong two-dimensionalism are modeled on Stalnaker’s diagonal propositions.
The identification of contexts and, more broadly, propositions with possible worlds is critical to the view outlined here. This identification, motivated, according to Soames, by Stalnaker’s view of conversation, allows for a very simple explanation of what it means to accept a proposition. It is to believe that the world with respect to which one is evaluating a proposition is a member of the set of worlds which constitute that proposition. This undergirds the legitimacy of taking the primary intension of S to assert the unrelativized truth of the character of S by explaining this assertion as the claim that the character of S is true with respect to some possible world W taken as a context of utterance iff the proposition assigned by the character of S at W contains W as a member.

By contrast, the weak two-dimensionalist simply identifies primary intensions with Kaplan-style characters and does not identify propositions with sets of possible worlds. Both views take the secondary intension of S at a world taken as a context W to be the proposition assigned by the character of S to W. Secondary intensions are typically, though not always, variable, as they are contextually determined.

For the strong two-dimensionalist, one understands a sentence S if and only if one knows its character and primary intension. For the weak two-dimensionalist, one understands S if and only if one knows its character (which is its primary intension, according to the weak theory). Such knowledge is all these views require, respectively, in order for a speaker to use S appropriately in a given context, and this knowledge in addition to total knowledge of the context always implies knowledge of the secondary intension of S in that context, but any ignorance of the context may result in a speaker’s not knowing the secondary intension expressed. Because it is the primary intension which is known by virtue of understanding the sentence, it is this
proposition which features in epistemic contexts, serves as the object of intentional attitudes, and is the bearer of a priority or a posteriority.

On both accounts, the secondary proposition is the bearer of modal properties. A secondary proposition is necessary if and only if it is true with respect to all world-states which are possible with respect to the world state of the context in which it is expressed. It is contingent if and only if it is false at some world-state which is possible with respect to the world-state of the context.

On the strong two-dimensionalist account, a sentence expresses a necessary a posteriori truth iff the secondary intension it expresses in a given context is true with respect to all worlds possible relative to that context and its primary intension is contingent, which is to say that the character of S assigns a false proposition to some context. A similar explanation is given for contingent a priori truths. A sentence S is a contingent a priori truth iff the secondary intension it expresses in a given context is true but not with respect to all worlds possible relative to the context and its primary intension assigns to all contexts a proposition which is true in that context. Such sentences are knowable a priori because one can know that the primary intension of S is true in a context without having any empirical knowledge of that context because it is true in all contexts. It follows that, on the strong two-dimensionalist account, no single proposition is both necessary and knowable only a posteriori or contingent and knowable a priori because the propositions which are the bearers of epistemic properties are primary intensions, and the propositions which are the bearers of modal properties are secondary intensions.

Weak two-dimensionalism contends that a sentence expresses a necessary a posteriori truth if its secondary intension is true with respect to all worlds possible relative to the world of the context, and its primary intension (i.e. character) cannot be known a priori to express a truth
in the context. Similarly, a sentence expresses a contingent a priori truth if its secondary
intension is true in a given context but false relative to some world which is possible relative to
the context, and its primary intension can be known a priori to assign a truth to the context.

It should be noted that the weak two-dimensionalist, unlike his stronger counterpart,
ads the possibility of a single proposition’s being both necessary and knowable only a
posteriori or contingent and knowable a priori. This is so because, for the weak two-
dimensionalist, to know the truth of a sentence is to know that its primary intension expresses a
truth in the relevant context.

On both two-dimensionalist accounts, proper names have their reference fixed by
rigidified descriptions and are in fact synonymous with these descriptions. Why do two-
dimensionalists insist on analyzing names descriptively? Let’s perform a two-dimensional
analysis of the sentence ‘Richard Nixon resigned the office of the Presidency.’ We know
perfectly well the secondary proposition expressed in the actual world taken as the context of
utterance. But what happens when we vary the context? If we try to imagine what proposition the
primary intension assigns to some context in which Richard Nixon never entered public life, we
face confusion as to the reference of the name ‘Richard Nixon.’ If we take the man who was
named Richard Nixon in the actual world as the referent of the name, are we not in some sense
considering the actual world as the context of the utterance?

Direct reference theories implicitly use the actual world as a constant context of
utterance. On a Kripkean account of reference, a speaker must refer to the actual world in order
to fix the reference of a name, but this is, in some sense, to treat it as a context. The two-
dimensionalist’s insistence on treating both contexts and circumstances of evaluation as variable
requires some manner of fixing the reference of names that is not dependent on invariably taking
a specific world as context. Given this requirement, descriptions seem the only suitable way of fixing reference. Although two-dimensionalism requires a descriptive analysis of names, it should be noted that the motivations for these views actually run in the other direction. Two-dimensionalism was developed in an attempt to revive descriptivism, but if descriptivism can be shown false, two-dimensionalism is indefensible. The two views stand or fall together.
Chapter 7:
Against Strong Two-Dimensionalism

There are a number of criticisms which apply to both the strong and weak versions of the two-dimensionalist view. In this chapter I will refrain from making those arguments and focus only on criticisms which apply exclusively to strong two-dimensionalism.

The claim that distinct propositions expressed by the same sentence serve as arguments under, respectively, modal and intentional operators is essential to strong two-dimensionalism. In many circumstances, everyday intentional attitudes and counterfactual claims can in fact be analyzed this way, but where modal and intentional operators interact, this claim is much harder to maintain.

Soames makes an argument which gets at this failure of integrating intentional and modal operators. Take the sentence

S₁ ‘The actual inventor of bifocals was actually the first Postmaster General of the United States’

For an agent to bear some intentional attitude to S₁ is for that agent to stand in a certain relation to the primary intension of S₁. For an agent to believe S₁ is to believe a proposition which asserts that the character of S₁ expresses a truth. Because the character of S₁ expresses a truth in a context iff the character of

S₂ ‘The inventor of bifocals was the first Postmaster General of the United States’

expresses a truth in that context, S₁ and S₂ have identical primary intensions. Thus to believe S₁ is to believe S₂ and vice versa. What is asserted by these belief ascriptions is that the same agent bears the same relation to the same thing (the identical primary intension of S₁ and S₂).
Given this equivalence, the two claims ‘it is a necessary truth that [if the actual inventor of bifocals was actually the first Postmaster General of the United States and Nathalie believes that the actual inventor of bifocals was actually the first Postmaster General of the United States, then Nathalie believes something true]’ and ‘it is a necessary truth that [if the actual inventor of bifocals was actually the first Postmaster General of the United States and Nathalie believes that the inventor of bifocals was the first Postmaster General of the United States, then Nathalie believes something true]’ are themselves equivalent. But the first claim is true while the second claim is false because if Nathalie were to believe $S_2$ in some context in which, say, John Quincy Adams had invented bifocals but Benjamin Franklin was the first Postmaster General, she would believe something false. Because the strong two-dimensionalist identifies belief in $S_1$ and belief in $S_2$, it seems he has no way of explaining the differing truth-value of these modal propositions which are, on his account, identical.

Soames gives a similar argument targeting the strong two-dimensionalist’s analysis of proper names as rigidified descriptions. Suppose that a particular object $O$, the Empire State Building, is denoted by the description $D$, ‘the most famous building in Manhattan.’ Now suppose that $N$, ‘The Empire State Building,’ is synonymous with a rigidified version of $D$, say ‘the actually most famous building in Manhattan.’ Once again, ‘Owen believes that $N$ is the $D$’ and ‘Owen believes that $N$ is the actual $D$’ are equivalent because the primary intensions of $D$ and ‘the actual $D$’ are identical. The strong two-dimensionalist is thus compelled to maintain that the two sentences

$$S_3 \text{ ‘Although Owen believes that } N \text{ is } D, \text{ had } N \text{ not been } D, \text{ Owen would not have believed that } N \text{ was } D’$$
S₄ ‘Although Owen believes that the D is D, had the actual D not been D, Owen would not have believed that the D was D’

are equivalent. But Owen might easily believe, in some world in which the most famous building in Manhattan were not the Empire State Building, that the most famous building in Manhattan was the most famous building in Manhattan. So S₃ might easily be true and S₄ false. The strong two-dimensionalist cannot explain this.

An analogous argument can be constructed using the operator ‘knows a priori.’ Take the sentences

S₅ ‘If there is a unique thing which is D, then the actual D is D’

S₆ ‘If there is a unique thing which is D, then the D is D’

The relation reported by ‘Stephen knows a priori that S₅’ is identical to the relation reported by ‘Stephen knows a priori that S₆’ because the ‘knows a priori’ operator takes the primary intension of each sentence as argument, and S₅ and S₆ have identical primary intensions. So again the claims ‘it is a necessary truth that [if Stephen knows a priori that if there is a unique thing which is D, then the D is D, then Stephen knows a priori that if there is a unique thing which is D then the D is D]’ and ‘it is a necessary truth that [if Stephen knows a priori that if there is a unique thing which is D then the D is D, then Stephen knows a priori that if there is a unique thing which is D, then the actual D is D]’ are equivalent on the strong two-dimensionalist account. But the latter claim is false because of course the actual D might not have been D, while the former is indeed true.

What these arguments show is that intentional and modal operators, when applied to a single sentence, take the same proposition as argument. This directly contradicts a central tenet of strong two-dimensionalism. Analogous arguments could easily be constructed using all
manner of intentional attitudes and epistemic operators, including ‘knows a posteriori,’ ‘desires,’ ‘believes,’ etc. The second argument also shows that descriptive analyses of names are not compatible with central strong two-dimensionalist beliefs about the functioning of primary and secondary intensions under modal and intentional operators.

Part of strong two-dimensionalism’s appeal stems from the pleasing symmetry between primary and secondary intensions when these are both construed as propositions. Moreover, propositions seem like the appropriate sort of thing to serve as arguments of both modal and intentional operators. It is obvious how secondary intensions can be construed as propositions, but it is not so clear that a proposition can play the role of context. For non-indexical sentences, no problems seem to arise, as everything relevant to the context can be gleaned from a total description of a world-state and propositions can be construed as sets of possible worlds. But sentences containing indexicals do pose problems. More is required than a total description of a possible world for something to serve as the context of an indexical sentence. Sentences containing ‘I’ and ‘now,’ for instance, require for interpretation knowledge of the speaker and time of utterance.

Chalmers suggests the notion of “centered” possible worlds as a way of resolving this difficulty:

There is a slight asymmetry in that a context of utterance but not a circumstance of evaluation is what Quine (1969) calls a centered possible world. This is an ordered pair consisting of a world and a center representing the viewpoint within that world of an agent using the term in question: the center consists in (at least) a “marked” individual and a time (1996, 60-1).
At first glance, this suggestion seems innocuous. It is more troubling when one realizes that these centered worlds are supposed to play a role in specifying the arguments of intentional operators. When the standard strong two-dimensionalist takes propositions to be sets of possible worlds, it is easy to see how primary intensions could be propositions. If they assert that the character of a sentence expresses a truth, then a given primary intension is the set of all possible worlds at which the character of its associated sentence assigns a proposition which is true with respect to those worlds taken as contexts. This becomes murkier when contexts are centered. A primary intension would then be the set of all centered worlds to which the character of its associated sentence assigns a proposition which is true. Chalmers calls such a primary intension a “centered proposition.”

It is far from clear that a set of centered possible worlds can reasonably be construed as a proposition. One might suppose that propositions are defined by two key semantic roles: bearing truth value and serving as the objects of intentional attitude ascriptions. It is easy to make sense of how propositions play these roles if one takes them to be sets of possible worlds. For a proposition (i.e. set of worlds) to be true when evaluated with respect to some world is for it to contain that world as a member. To, for example, believe a proposition with respect to some world is to believe that it contains the relevant world as a member. This analysis seems to break down when the worlds in question are centered.

If contexts are identified with centered possible worlds, then to believe some sentence is to believe that its primary intension expresses a truth at some particular centered world. At this point, there are two approaches a two-dimensionalist might take to explain how centered propositions are to be evaluated. One might argue that centering a world does not add any fact to that world, so a centered version of a particular world is identical with its uncentered counterpart.
If one accepts this, believing a proposition with respect to some world $W$ could be analyzed as believing that some centering of $W$ is a member of the set of centered worlds constituting the proposition in question. Thus one could maintain that the circumstances with respect to which centered propositions are to be evaluated are uncentered worlds. A centered proposition would then be true with respect to some uncentered world if some centering of that world (the circumstance of evaluation) is a member of the set of centered worlds constituting the proposition to be evaluated.

This analysis gives a very misleading account of what is believed when I assert, for example, that ‘I believe I am the President of the United States’ when evaluated with respect to the actual world. Of course, there is a centering of the actual world which belongs to the set of centered worlds constituting this “proposition,” but I seem to assert a rather more grand belief than that. This example also belies the identification of centered and uncentered versions of the same world, for Barack Obama and I might share the belief reported above, thus believing the same centered “proposition,” but when evaluated with respect to the actual world, he believes something true and I believe something false. This approach is irredeemably flawed.

The alternative is to evaluate centered propositions with respect to centered worlds. On this approach, the belief expressed by ‘I am the President of the United States’ is true only when evaluated with respect to a world centered on the President of the United States. This seems to give the correct result, but it makes the truth of a proposition relative to a particular viewpoint. One cannot say whether a proposition is true with respect to a particular world $simpliciter$, but only with respect to a particular viewpoint in that world. Insofar as this result is troubling, it should cast doubt on this strategy.
The problems raised by indexicals run even deeper, as a number of Soames’ arguments show. Take the indexical sentences

\[ S_a \] ‘I am here now’

\[ S_b \] ‘I am not Bernie Sanders’

\[ S_c \] ‘I know that I am here now’

\[ S_d \] ‘I know that I am not Bernie Sanders’

When I assert \( S_c \) or \( S_d \) I presumably report knowledge of the same propositions that I report Sanford Shieh as knowing when I assert

\[ S_e \] ‘Sanford Shieh knows that I am here now’ and

\[ S_f \] ‘Sanford Shieh knows that I am not Bernie Sanders’

Moreover, it seems these are the very same propositions I report Taylor Swift as not knowing when I assert

\[ S_g \] ‘Taylor Swift does not know that I am here now’ and

\[ S_h \] ‘Taylor Swift does not know that I am not Bernie Sanders.’

It seems in all these cases that we are dealing with the same two propositions to which various agents stand in various epistemic relations, but it is not clear on the two-dimensionalist account which two propositions these are. One would expect the primary intensions of \( S_a \) and \( S_b \) to play all these various roles, but clearly they cannot because each agent would accept these intensions as expressing a truth in his or her own context. Moreover, my knowledge is a priori whereas Sanford Shieh can only know these truths by recourse to empirical evidence.

Soames discusses a few potential two-dimensionalist counterarguments to the problems raised by indexical sentences. One might argue that sentences containing indexicals in the complement clauses of attitude ascriptions ought to be analyzed by replacing such indexicals
with variables and assessing their truth value only relative to an assignment of some argument. So \( S_f \) would be analyzed ‘Sanford Shieh knows that \( x \) is not Bernie Sanders’ and determined as true with respect to an assignment of me to \( x \) with the actual world as context. The character of the expression ‘\( x \) is not Bernie Sanders’ relative to an assignment of me to \( x \) is a constant function from any possible world taken as a context of utterance to the same proposition because names have stable character. The primary intension of this expression would then be something like a proposition asserting that this character assigns a proposition which is true in a way unrelativized to a context, which isn’t particularly salient here anyway. It is hard to even formulate what the primary intension of such an expression would be, which seems to hint at how useless and ill-defined the notion is in this case. It is hard to see how the primary intension of the sentence, whatever it is exactly, is what Sanford Shieh knows.

What Sanford Shieh knows is a particular fact about me, that I am not Bernie Sanders, which is the secondary intension, in the actual context, of the relevant sentence. Moreover, nothing about the suggested analysis changes the necessary a posteriori status of the proposition Sanford Shieh is reported to know. The proposition expressed by ‘\( x \) is not Bernie Sanders’ relative to an assignment of me to \( x \) is a necessary truth, and yet Sanford Shieh can know this truth only a posteriori. The suggested analysis fails to resolve any of the problems raised for strong two-dimensionalism by the previous example.

There is one other approach which a strong two-dimensionalist might take here. Suppose we assume that indexicals in the complement clauses of attitude ascriptions are to be replaced by co-referential rigidified descriptions. On this approach, \( S_f \) is analyzed, after replacing ‘Bernie Sanders’ with a synonymous rigidified description, as ‘Sanford Shieh knows that dthat [the \( x: x \) is Sam McNutt] is not dthat [the \( x: x \) is Bernie Sanders].’ Given that the dthat-operator only
serves to rigidly fix a reference relative to a particular context, it can be dropped in discussing the primary intension of the complement clause because the primary intension is not relativized to any particular context. So the primary intension which Sanford Shieh is reported as knowing is that which is associated with ‘[the x: x is Sam McNutt] is not [the x: x is Bernie Sanders].’ This contingent proposition is, on the view described, what is known (a posteriori). Thus the strong two-dimensionalist theses that distinct propositions serve as arguments of, respectively, intentional and modal operators and that no single proposition is both necessary and knowable only a posteriori can be maintained.

Despite its apparent success, this approach reveals deep flaws upon examination of the descriptions employed. I examined various approaches to rigidifying descriptions, and it seemed the only plausible candidates made use of Kripke’s suggestion that reference is fixed by some sort of causal-historical chain linking the speaker to the initial baptism of the name. Let us exemplify these by a description suggested by David Lewis, ‘the individual whom I have heard of under the name ‘n’.’ When one realizes that this is the sort of description which the strong two-dimensionalist employs in the above analysis of indexical sentences, the problem is obvious. One indexical has simply been traded for another.

Let us undertake our earlier analysis again, this time employing an actual description of the relevant kind. What Sanford Shieh knows is supposedly that ‘dthat[the x: x is Sam McNutt] is not dthat[the individual I have heard of under the name ‘Bernie Sanders’].’ We have now reintroduced an indexical into the complement clause of an attitude ascription and all the problems thus entailed. If we perform the suggested procedure for analyzing the indexical again, we get ‘dthat[the x: x is Sam McNutt] is not dthat[the individual whom dthat[the x: x is Sam McNutt has heard of under the name ‘Bernie Sanders’]]. What sort of description might fill the
role of ‘the x: x is Sam McNutt,’ without running afoul of Kripke’s semantic and epistemic arguments? Certainly not one containing a proper name, for it would in turn have to be analyzed as a description. The only plausible candidates contain indexical terms, and all such descriptions run into the problems facing ‘the individual whom I have heard of under the name ‘n.’’ There simply is no conceivable description which is capable of serving this purpose.

There is an additional problem which Soames points out. Sanford Shieh’s knowledge does not depend on whom I have heard of under the name ‘Bernie Sanders,’ so the analysis is empirically false. One might here propose that ‘Bernie Sanders’ ought instead be analyzed ‘dthat [the individual whom Sanford Shieh has heard of under the name ‘Bernie Sanders’].’ But this is, after all, my assertion, and it does not seem as though its content should be made dependent on what Sanford Shieh has or has not heard. In this case, I would not even know what I was asserting, let alone whether or not it was true.

These arguments demonstrate that, for sentences containing names or indexical terms that express necessary a posteriori or contingent a priori truths, descriptions must be found which can substitute for the relevant names. The only descriptions which seem plausible candidates themselves contain names or indexical terms, which must then be explained away. It is impossible ever to reach the end of this regress. There are also irresolvable problems that arise from trying to construe sets of centered worlds as the objects of intentional operators. The claim that such operators take different objects than modal operators when applied to the same sentence is untenable. Not only is the appeal of strong two-dimensionalism undermined by Soames’ numerous arguments, the position is demonstrably false when confronted with basic intuitions about intentional attitudes and modality.
The arguments in the previous chapter apply only to the strong version of the two-dimensionalist theory. I will now examine a number of Soames’ arguments against weak two-dimensionalism.

Let’s rehash the basic distinctions between the strong and weak versions of the theory. Whereas the strong two-dimensionalist takes the primary intension of a sentence to be a proposition associated with its character, the weak two-dimensionalist simply identifies primary intensions with characters. The weak two-dimensionalist eschews the identification of propositions with sets of possible worlds. The weak two-dimensionalist accepts that a proposition can be both necessary and knowable only a posteriori or contingent and knowable a priori and that secondary intensions bear both epistemic and modal properties.

On the weak two-dimensionalist account, an agent bears an intentional attitude to a proposition (secondary intension) $p$ expressed by a sentence $S$ in virtue of the relation in which he stands to the primary intension (character) of $S$. For example, to know $p$ in a particular context is to know that the primary intension of $S$ expresses a truth in that context. Given this analysis of intentional attitude ascriptions, the weak two-dimensionalist can formulate deflationary accounts of the necessary a posteriori and contingent a priori: a sentence $S$ expresses a necessary a posteriori truth in $C$ iff its secondary intension is necessary and knowable in virtue of accepting that its associated primary intension expresses a truth in $C$, this primary intension assigns the relevant secondary intension to $C$ but assigns a false proposition to some other context, and knowing this primary intension to express a truth in $C$ is only possible by recourse to empirical evidence. A sentence $S$ expresses a contingent a priori truth in $C$ iff its secondary
intension is contingently true and knowable in virtue of accepting that its associated primary intension expresses a truth in C, this primary intension assigns the relevant secondary intension to C and assigns true propositions to every other context, and knowing this primary intension to express a truth in C does not require empirical evidence.

This account is deflationary because the a posteriority and a priority involved, although properties of secondary intensions, result from the relation in which an agent stands to a primary intension. Instead of a full-blooded assertion that a single true proposition can be both metaphysically necessary and knowable only by recourse to empirical evidence, the weak two-dimensionalist asserts that empirical evidence may be required to know that the character of a sentence expresses a truth, which may turn out to be necessary, in a particular context.

Soames raises two possibilities which a weak two-dimensionalist cannot classify in the most intuitively straightforward ways:

(i) There is a necessary proposition p expressed by a sentence S satisfying the following conditions: (a) the character M of S expresses a truth in every context, (b) one can know p by virtue of understanding S, justifiably accepting it (i.e., accepting M), and knowing it to express a truth, (c) justifiably accepting M and knowing it to express a truth requires empirical evidence, and (d) there is no other, a priori, route to p.

(ii) There is a contingent proposition p expressed (in some context C) by a sentence S satisfying the following conditions: (a) the character M of S expresses a truth in C, but in other contexts it expresses something false, (b) in C, one can know p by virtue of understanding S, justifiably accepting it (i.e., accepting M) and knowing it to express a truth (in C), and (c) justifiably accepting M and knowing it to express a truth (in C) does not require empirical evidence (294).
On the weak two-dimensionalist account, (i) is not an instance of the necessary a posteriori, for a proposition is knowable only a posteriori iff its character assigns a false proposition to some context. The weak two-dimensionalist cannot classify (i) as a priori because he maintains that a priori knowledge never requires recourse to empirical evidence. He cannot classify (ii) as a priori because, on the two-dimensionalist account, a proposition can only be known a priori if its primary intension can be known to express a truth in every context, but he cannot classify (ii) as a posteriori because he is committed to the claim that no proposition is knowable only a posteriori if it can be known simply by justifiably accepting that the character of the sentence which expresses p expresses a truth in the relevant context (without resort to empirical evidence). Because the weak two-dimensionalist has no way to classify (i) and (ii), it seems he must deny that there are such propositions.

The strong two-dimensionalist who identifies propositions with sets of possible worlds has no difficulty explaining why there can be no propositions such as (i) and (ii). If a proposition is necessary, it is the set of all possible worlds. No empirical evidence could possibly be needed to determine that a necessary proposition was true because any world at which it could be evaluated would be a member of the set of all possible worlds. Similarly, a proposition such as (ii) is impossible because a contingent proposition is the set of worlds at which the primary intension of the associated sentence expresses a truth. There could be no way other than empirical inquiry to determine whether or not the world with respect to which one is evaluating (ii) is a member of that set.

The weak two-dimensionalist does not have recourse to such explanations. Though he is committed to denying the possibility of propositions such as (i) and (ii), he has no rationale for doing so. It is worth pointing out how easily we might come by such propositions. A sentence in
some possible language which rigidly designates an object by name and asserts a necessary property of that object would be an instance of type (i). It is more doubtful whether propositions of type (ii) are possible, and in fact I think they are not. Nevertheless, neither possibility is wildly implausible, and it seems quite likely that there are type (i) propositions. Some argument is certainly required from those who would deny such the possibility of there being such propositions, and it is doubtful the weak two-dimensionalist can muster one.

The weak two-dimensionalist is committed to the claim that all names have their reference fixed by some description(s) which does not itself contain any names. This is typically claimed to be a priori, on grounds such as Frank Jackson describes in “Reference and Description Revisited”:

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 (212).

Jackson thinks names and natural kind terms must have their reference fixed descriptively simply because it is possible to answer questions about how such terms would refer in described counterfactual circumstances.

There are a number of reasons why such a bold defense of descriptivism fails. First, there are circumstances in which there is no difficulty determining the reference of terms for which no unique reference-fixing description is possible. David Kaplan suggests an example of such a circumstance. Suppose two twins, Castor and Pollux, are raised in identical environments so as to be identical even at the molecular level. It would seem that such twins could not but associate identical descriptions with any linguistic terms they use, yet each refers unproblematically to himself when he uses ‘I.’

Even more damaging to Jackson’s argument is that he confuses facts about the meaning of terms with facts about what is required in order for terms to have the meaning and reference that they do. Descriptivism requires that descriptive knowledge be a component of meaning. Jackson’s argument from the possibility of determining reference in various described scenarios at best implies that descriptive knowledge is prerequisite for using some names and natural kind terms. It does not follow from the claim that we have some descriptive knowledge which enables us to determine the reference of the term ‘water’ in various described scenarios that this knowledge is part of the meaning of ‘water.’

Even Jackson’s premise is doubtful. People are thoroughly familiar with language and use it constantly. It is not surprising that they have nuanced views about how various linguistic
expressions would refer in various situations which are described to them without attributing to them any conscious or unconscious theory of reference.

It seems Jackson is far too bold in his defense of descriptivism. He fails to show that names and natural kind terms must have their reference fixed by description. Any argument for descriptivism must make a positive case for why reference is fixed in that particular way, as there are plausible alternatives. A number of these arguments are rebutted in chapter 4, including attempts to make use of the causal-historical account of reference fixing and the use of the actuality and dthat operators in rigidification. The weak two-dimensionalist commitments that names have their reference fixed descriptively and are synonymous with rigidified descriptions cannot be maintained.

The synonymy of names and natural kind terms with rigidified descriptions was the basis for the two-dimensionalist’s treatment of such terms on the model of Kaplan’s treatment of indexicals. Since these non-descriptive terms don’t share the context-variability of indexicals, their characters are constant functions. Given the weak two-dimensionalist identification of primary intensions with characters, their primary intensions simply map the same secondary intension to each circumstance of evaluation, and thus the weak two-dimensionalist analysis of the necessary a posteriori and contingent a priori falls apart. The demonstrable failure of descriptivism dooms the weak two-dimensionalism which was meant to save it.
Interest in linguistic reference is understandable, indeed unavoidable, wherever there is concern for truth. In order to know whether we have got hold of a truth and what sort of truth it is, we must know something about the language in which it is expressed. The reference of proper names may seem an inconsequential matter, peripheral to these weighty issues, but names are unusual words, and philosophers have always been well served by attending to the exceptions, to things which don’t fit neatly.

Descriptivism is an attempt to assimilate proper names into linguistic categories whose means of referring is more readily understood. Kripke vindicated our common intuition of the uniqueness of names by outlining the special ways by which they acquire and retain reference. By drawing out the implications of names’ unusual referential properties, intriguing modal and epistemic claims about certain propositions containing names can be deduced.

Kripke’s arguments for the direct reference of names and natural kind terms threatened many cherished philosophical commitments. A backlash was inspired, and an attempt to revive descriptivism began. A hunt for reference-fixing descriptions which could evade Kripke’s semantic argument turned up a few plausible candidates, all of which attempt somehow to capture the causal-historical source of reference descriptively. However, all of these candidates ultimately fail to properly perform their intended function in certain cases. I cannot say it is impossible to find descriptions which fit the bill, but it looks that way. Attempts were then made to rigidify reference-fixing descriptions so that they might evade Kripke’s modal argument. It is possible to do this with Kaplan’s dthat-operator, but in the process much of descriptivism’s appeal is lost. Semantic two-dimensionalism is the final step in this attempted revival.
The first step in constructing a two-dimensionalist semantic framework is to associate distinct primary and secondary intensions with each sentence. Much of this is actually an unobjectionable extension of Kaplan’s insights into the distinction between character and content. Things go wrong when the strong two-dimensionalist tries to construe primary intensions as propositions on the model of Stalnaker, whose notion of a diagonal proposition is itself incoherent. The confusion only mounts when the strong two-dimensionalist tries to construct a primary intension from this muddled material. The incoherence in the strong two-dimensionalist notion of a primary intension is one of the view’s most damning features. It is difficult to lay out a clean argument against the identification of primary intensions with propositions asserting that the character of their associated sentences expresses a truth unrelativized to any particular context because it is impossible to pin down what is meant by this. Nevertheless, sufficient reflection on what this could possibly mean ought to convince one to reject the notion.

The strong two-dimensionalist claim that distinct propositions associated with a single sentence serve, respectively, as arguments of modal and intentional operators is false. Common sense analyses of sentences in which modal and intentional operators interact contradict this claim, and no two-dimensionalist explanation seems possible.

Sentences containing indexical terms also confound strong two-dimensionalist analysis. Attempts to avoid these problems by treating possible contexts of utterance as centered worlds threatens the identification of propositions with sets of possible worlds and runs the risk of making the truth of a proposition relative to a particular point of view. The problems posed by indexicals are particularly troubling for two-dimensionalism given its requirement that names be
analyzed as suitable rigidified descriptions, all of which seem to contain indexicals. For all of these reasons, strong two-dimensionalism is false.

Weak two-dimensionalism avoids most of the problems facing the strong variant of the theory. Soames raises some examples of sentences which the weak two-dimensionalist cannot classify and the possibility of whose existence he must deny. This should cast serious doubt on the theory, as Soames’ examples seem at least plausible and not to be denied out of hand. What really condemns weak two-dimensionalism is the impossibility of identifying names with rigidified descriptions which was the basis for treating them on Kaplan’s model of indexicals. If names have constant characters, there is no sense in giving them a two-dimensional analysis.

Again, descriptivism and two-dimensionalism stand or fall together. Strong descriptivism, which identifies the meaning of a name with some description(s), is demonstrably false, and it is this version of the description theory which two-dimensionalism requires. On these grounds, weak two-dimensionalism is false.

Both strong and weak versions of two-dimensionalism are riddled with errors and confusions. The only sort of description theory which might be defensible is a weak reference-fixing version which employs Kaplan’s dthat-operator. Even then, the difficulties involved in finding descriptions immune from Kripke’s semantic argument are immense and quite possibly insurmountable. In the face of so many forceful and intuitively appealing arguments for a direct reference theory of names and natural kind terms, it is time to abandon descriptivism.

Debates over two-dimensionalism have only prolonged descriptivism’s inevitable demise. The debate is over. Two-dimensionalism is unworkable. Descriptivism is dead. It is time to accept the basic tenets of the Kripkean account of reference and move on to new linguistic problems.
Works Consulted


