Margolis, E., & Laurence, S. (2007). The Ontology of Concepts: Abstract Objects or Mental Representations?. *Noûs*, 41:4, pp. 561-593. DOI: 10.1111/j.1468-0068.2007.00663.x

This is a preprint. For the final version, please visit Noûs.

The Ontology of Concepts—Abstract Objects or Mental Representations?

Eric Margolis and Stephen Laurence [1]

> *The word "concept" is used in various ways; its sense is sometimes psychological, sometimes logical, and sometimes perhaps a confused mixture of both.*
> (Gottlob Frege 1892a, p. 42)

What is a concept? Philosophers have given many different answers to this question, reflecting a wide variety of approaches to the study of mind and language. Nonetheless, at the most general level, there are two dominant frameworks in contemporary philosophy. One proposes that concepts are mental representations, while the other proposes that they are abstract objects. This paper looks at the differences between these two approaches, the prospects for combining them, and the issues that are involved in the dispute. We argue that powerful motivations have been offered in support of both frameworks. This suggests the possibility of combining the two. Unlike Frege, we hold that the resulting position is perfectly coherent and well worth considering. Nonetheless, we argue that it should be rejected along with the view that concepts are abstract objects—the two have a shared failing—and that the mental representation framework is to be preferred.

Here is how we proceed. In sections 1 and 2, we introduce the two frameworks and briefly review their supporting motivations. Next, in section 3 we show how mental representations and

[1] This paper was fully collaborative; the order of the authors' names is arbitrary.

abstract objects can be combined into a unified theory of concepts—what we call the *Mixed View*. Much of the rest of the paper is then concerned with objections to the Mixed View. In sections 4-7, we examine objections that focus on the Mixed View's commitment to mental representations. These include the threat of an explanatory regress, the claim that the view implicates the wrong kinds of mental representations (structured ones), the charge that it lacks sufficient generality, and a variety of concerns associated with neo-Behaviorist approaches to the mind. We argue that each of these objections is flawed and that the commitment to mental representations is perfectly sound. In section 8, we go on to argue that the situation isn't as cheerful for the Mixed View's abstract objects. These sense-like entities are problematic because they are unable to fulfill their explanatory function of solving the mode of presentation problem. The final question, then, is whether mental representations by themselves can solve the mode of presentation problem. In section 9, we explore a number of promising strategies that suggest they can.

1. Concepts as Mental Representations

The view that concepts are mental representations takes as its starting point a version of the *Representational Theory of the Mind* (RTM). According to RTM, thinking occurs in an internal system of representation; occurrent propositional attitudes are token mental representations (i.e., mental particulars with semantic properties). For example, when someone arrives at the belief that her house needs a new coat of paint, RTM says that she comes to form a mental representation, one that represents her house and its state of disrepair. What makes this representation a belief is its functional role; it figures in mental processes in the way that is characteristic of beliefs. Desires and other attitude types have their own distinctive functional roles.[2] The version of RTM that is relevant to our discussion maintains further that the internal system of representation has a language-like syntax and a compositional semantics and, to that extent, amounts to a language of thought (Fodor 1975). The analogy with natural language

[2] RTM is standardly presented as taking beliefs and other propositional attitudes to be *relations* between agents and mental representations. But given that the relation in question is a matter of an agent's having a token representation with a particular type of functional role, it is simpler to say that beliefs just are *mental representations with a characteristic type of functional role*.

syntax is meant to indicate that the internal system of representation incorporates, at the very least, a distinction between predicates and subjects, and that it includes logical devices, such as quantifiers and variables. That the system has a compositional semantics is just to say that the content of a complex representation is a function of its syntax and the contents of the representations from which it is composed. On a view like this, the belief that one's house needs a new coat of paint is a structured psychological entity. Among its constituents are representations about the house and paint.

For anyone who embraces this framework, it's natural to maintain that concepts are mental representations, in particular, the representations out of which propositional attitudes are composed. We call this the Psychological View of concepts:

The Psychological View of Concepts

Concepts are mental representations. They are the constituents of propositional attitudes such as beliefs and desires.

The Psychological View is the default position in many areas of cognitive science and enjoys a good deal of support in the philosophy of mind. It is at the center of a rich and powerful model of the mind, but two of its benefits are especially worth mentioning.

The first is the promise of explaining the productivity of thought. *Productivity* refers to the fact that, under suitable idealization, there is no upper bound to the range of semantically distinct thoughts. One way of appreciating just how vast our cognitive capacities are is to consider the thoughts associated with the sentences of a natural language. As Noam Chomsky (1972) has noted, nearly every sentence we speak or hear is a sentence we have never before encountered, but despite the novelty of these sentences, we have no difficulty entertaining the corresponding thoughts. (The sentences of this paper are an example. It's unlikely that readers have come across most of these very sentences before.) The psychologist George Miller (1995) makes the point all the more vivid by focusing on just 20-word sentences, asking how many these we can understand. Assuming conservatively that there are on average 10 words to draw from for each word choice as a sentence is constructed, the implication is that we understand at least 10^{20} 20-word sentences. That's one hundred million trillion of them. By comparison, the human brain contains roughly 10^{11} neurons, and the number of seconds in the history of the Universe is

estimated to be on the order of 10^{17}. So assuming that each sentence corresponds to a distinct thought,[3] and sticking only to 20-word sentences (that is, ignoring not just longer sentences but also shorter ones), the number of thoughts we arrive at is more than a billion times the number of neurons in the brain and roughly a thousand times the number seconds in the history of the Universe.[4] According to RTM and the Psychological View, this is just the tip of the iceberg. Once we abstract away from limitations of memory and attention and other factors that interact with our thinking, the human capacity for entertaining new thoughts is without limits. The actual thoughts that people entertain in their lifetime constitute a tiny and idiosyncratic subset of the thoughts that their conceptual system makes possible.

RTM's explanation of productivity is in terms of the compositional rules that govern the domain of mental representations. A finite stock of basic concepts can be combined over and over again, using the same rules to create increasingly complex mental representations. The point is that without mental representations, and without their having compositional structure, a fundamental fact about minds like ours is left wholly unexplained. We'd have no account of how our finite resources manage to achieve an unbounded representational potential (Fodor 1987).[5]

The second explanatory benefit of the Psychological View and the version of RTM that it accompanies is that together the two promise to explain how mental processes can be rational and yet realized in a physical system—how thinking can be mechanized (Fodor 1987). The answer is given in terms of the dual life that structured mental representations possess. They have physical-causal (presumably neurological) properties but they also have semantic properties. The two are coordinated in much the same way that they are coordinated in digital computers. A digital computer is a physical device whose state transitions are arranged in such a way that its physical operations respect the interpretation of its symbols. By taking mental

---

[3] Some of the corresponding thoughts may arguably be the same (e.g., ones corresponding to active and passive constructions), but even if we don't make this simplifying assumption, it should be clear that the number of corresponding thoughts is astounding.

[4] Notice that 20-word sentences needn't seem particularly long or taxing. For example, the sentence this note is attached to is 61 words long. The preceding sentence is 29 words long.

[5] We are inclined to push this argument more forcefully than Fodor, who, in his recent writings, emphasizes systematicity over productivity. (Propositional attitudes are supposed to be systematic in that the ability to entertain one brings with it the ability to entertain others that are semantically related.) Fodor's reason for downplaying the productivity argument is that it requires an idealization that some may not be prepared to grant. But idealization is a ubiquitous feature of theorizing about natural phenomena, and we see no reason why it should be any more problematic in the study of the mind than elsewhere.

processes to be computations over internal representations, RTM opens the way for a parallel treatment of the rationality of human thought and action. It begins to explain how thinking can involve moving from one thought to another in a coherent manner that directs an agent's behavior in such a way as to fulfill her goals. Why did Sue reach for the glass of wine? Because she *wanted* some wine and she *believed* that the glass contained wine. But how do these thoughts—*Sue's wanting some wine, Sue's believing the glass contains wine*, etc.—translate into her actually obtaining the wine? RTM is able to explain this by its appeal to computations over structured representations that depict what Sue wants and the situation that she finds herself in. As a result, RTM shows how the causally efficacious states that guide our behavior can be identified with the reasons that are responsible for that behavior. This is no small payoff and, together with the account of productivity, it provides considerable motivation for adopting the Psychological View of concepts.

2. Concepts as Abstract Objects

Let's now turn to the second major view of the ontological status of concepts. According to this view, concepts aren't mental representations; they aren't "in the mind" at all. Instead, they are abstract objects of a certain sort. To understand this view, it helps to see how it emerges from the tradition in semantics stemming from the seminal work of Gottlob Frege.

The Fregean tradition maintains that the meaning of a declarative sentence is a proposition, where propositions are understood to be abstract objects that exist independently of our minds and that are the primary bearers of truth.[6] Sentences in different languages that mean the same thing express the same proposition, and understanding the meaning of a sentence involves grasping the proposition that it expresses.

Propositions are notoriously elusive entities. But the commitment to propositions is supposed to be vindicated by the explanatory work that propositions do in semantics. For example, belief ascriptions—sentences like (1)—are said to relate a believer to an entity:

[6] In Frege's system, propositions are called *thoughts*. This introduces considerable potential for terminological confusion. To keep things straight, it's important to remember that Fregean thoughts are not psychological entities at all and, in particular, they are not occurrent mental states. They are abstract entities of the sort described in the text.

(1) Jill believes that grass is green.

Why? Notice that it seems to follow from (1) that there is something that Jill believes. And if Bob also believes that grass is green, then it seems that there is something that Jill and Bob both believe, viz., *that grass is green*. What is this thing *that grass is green*? According to the Fregean tradition, it is a proposition. Numerically one and the same entity—the proposition *that grass is green*—is the object of both Jill and Bob's belief. It doesn't exist in a particular place (abstractness), it could have existed whether or not anyone ever believed it (mind-independence), and it is owing to the truth of the proposition *that grass is green* that the belief that Jill and Bob share is true (primary bearer of truth).

The Fregean tradition that we have in mind goes further in taking propositions to be structured entities (Peacocke 1992, Zalta 2001). Just as a sentence is composed of words and phrases, a proposition is supposed to be composed of more basic elements. Moreover, these elements are supposed to have the same ontological standing as propositions themselves. They too are abstract objects, where different tokens of the same belief involve numerically the same concepts.[7] The second major view of the ontology of concepts identifies concepts with the abstract constituents of Fregean propositions. We call this the Semantic View of Concepts.

The Semantic View of Concepts

Concepts are abstract objects; they are the constituents of Fregean propositions.

In Frege's framework, the constituents of propositions are called *senses* and are associated with a number of explanatory functions. At the very least, they contribute to a theory of what it is to know the meaning of a word. To know a word's meaning is to grasp its sense. In addition, senses are crucial to explaining how expressions can have a meaning despite lacking a reference (e.g., "Pegasus" isn't meaningless). Meaning without reference is possible because an empty term can still have a sense. Most importantly, senses are invoked to solve the *mode of presentation problem*, that is, the problem of explaining how it is that co-referring expressions

[7] See Sutton (2004) on how best to characterize views that treat concepts as abstract objects. We follow Sutton in taking the central point to be that on such views individual concepts don't admit of a type-token distinction. In contrast, the Psychological View fully embraces the idea that concepts have tokens (token concepts are token mental representations).

can have differing cognitive significance. For example, though the names "George Orwell" and "Eric Blair" refer to the very same person, "George Orwell wrote *1984*" needs to be distinguished from "Eric Blair wrote *1984*", since we might infer different things from them, and one could reasonably assent to the first while disavowing the second. Similarly, "George Orwell is George Orwell" smacks of triviality, while "George Orwell is Eric Blair" makes a substantive claim. For Frege, the explanation of these differences is that there is more to semantics than reference. Although "George Orwell" and "Eric Blair" refer to the same man, an understanding of these terms involves grasping their respective senses—senses that present the man in different ways, or as Frege sometimes puts it, that contain different *modes of presentation* of their referent.

The explanatory functions associated with modes of presentation are the primary concern for advocates of the Semantic View. For example, in locating the subject matter of a theory of concepts, Christopher Peacocke remarks that:

> Concepts *C* and *D* are distinct if and only if there are two complete propositional contents that differ at most in that one contains *C* substituted in one or more places for *D*, and one of which is potentially informative while the other is not. (Peacocke 1992, p. 2)

In other words, *C* and *D* are different concepts because they embody differing modes of presentation. Peacocke goes on to note that the sorts of things he calls "concepts" aren't what Frege called "concepts" but that the disagreement is purely terminological. Peacocke and other proponents of the Semantic View employ the term to refer to Fregean senses, whereas Frege uses it to pick out the referents of predicates.[8] To avoid any possibility of confusion, we should note here that we are following the contemporary usage (Peacocke's). According to the Semantic View, then, concepts are senses.[9]

Taking concepts to be senses singles out a different explanatory profile than we saw with the Psychological View. On the Semantic View, concepts are most directly associated with concerns

[8] For Frege, these referents are functions from objects to truth values. This may make them abstract objects as well, but they are still to be distinguished from senses in that functions are individuated extensionally—so if function *f* and function *g* have the same extension, then *f* = *g*.

[9] An anonymous referee has suggested that a different way to treat concepts as abstract objects would be to adopt Wayne Davis's view that concepts are mental representation *types* (Davis 2003, 2005). Although types are certainly abstract entities, the suggestion threatens to eliminate the fundamental difference between the Psychological View and the Semantic View, viz., the disagreement about whether concepts have tokens (see note 7). On our taxonomy, Davis's view turns out to be a version of the Psychological View, not the Semantic View.

about linguistic meaning and cognitive content. But proponents of the Semantic View have also motivated their approach by noting problems with mental representations. For example, Frege himself dismisses mental representations, claiming they are too subjective. He argues that communication requires shared meanings but that people can't literally have the same mental representations since mental representations are components of our individual subjective experiences. Senses, on the other hand, are objective and eminently shareable. Two people "are not prevented from grasping the same sense" (Frege 1892b, p. 60), as senses are external to our minds. Peacocke cites a related objection to mental representations. He claims that "[i]t is possible for one and the same concept to receive different mental representations in different individuals" (Peacocke 1992, p. 3). And Georges Rey (1994) adds that the very same mental representation can correspond to distinct thought contents, for instance, an image of Paris can correspond to thoughts of Paris but also to thoughts of France. Putting all this together, the prospects for mental representations may seem dim and the need for senses all the more pressing. Defenders of the Semantic View conclude that we should identify concepts with the constituents of Fregean propositions.

3. A Mixture of Both?

How do the Psychological View and the Semantic View relate to one another? Clearly, some authors see them as standing in opposition. But it's not unreasonable to ask whether the two could be combined in a way that capitalizes on the advantages that each has to offer.[10] In this section, we discuss the possibility of a mixed theory, one that incorporates both mental

[10] A different option is to say that there really is no substantive dispute here at all, only the terminological matter of which things to call 'concepts'. Perhaps we should just introduce two terms—*concept*$_1$ (for mental representations) and *concept*$_2$ (for senses)—and be done with it. While we can't entirely rule out this interpretation, we do think it would be a mistake to trivialize the differences between the Psychological View and the Semantic View. As we mention in the text, some of their proponents certainly don't see the matter as a merely terminological one; they argue quite vigorously against the opposing position. In addition, one ought to be hesitant about claiming that the dispute is merely terminological for the simple reason that *any* substantive dispute can be recast as a terminological one. For example, we can call consequentialist goodness *good*$_1$ and deontological goodness *good*$_2$, or compatibilitist freedom *freewill*$_1$ and incompatibilist freedom *freewill*$_2$, and then go on to pronounce that philosophers on different sides of the original disputes were just talking about different things. But such pronouncements clearly don't resovle the original controversies. Complex issues remain concerning whether we need either or both of the theoretical kinds identified and which on balance best answers the explanatory concerns that preceded the terminological innovations.

representations and senses. We argue that combining the two offers a perfectly coherent theoretical option. The first step towards achieving such a reconciliation is assessing the arguments that proponents of the Semantic View have employed against mental representations. As it turns out, these arguments aren't compelling.

Frege's own case against mental representations is especially disappointing. The initial rejection is based on the claim that people can't literally share mental representations because mental representations figure in conscious subjective experience and hence are unique to the minds in which they occur. One quibble with this argument is that Frege is clearly assuming that mental representations are conscious mental images, while contemporary theories of mental representation aren't restricted in this way. The main problem with the argument, however, is that its plausibility turns on a type-token confusion. Frege is certainly right that each person has her own unique mental representation tokens, but the question is whether different tokens in different minds can be of the same type, and we see no reason why they can't be. The situation is, in principle, no different than it would be had Frege argued that two people can't literally utter the same sentence. While it's true that each will produce her own token, that doesn't mean that the utterances can't be instances of the same sentence *type*, for example, both instances of the sentence "snow is white".

Frege seems to concede this point only to dismiss its significance. He says, "It is indeed sometimes possible to establish differences in the ideas [mental representations] … of different men; but an exact comparison is not possible, because we cannot have both ideas together in the same consciousness" (Frege 1892b, p. 60). The argument, in other words, is that we can never really be sure that other people associate the same mental representations (i.e., representations of the same type) with the same words as we do. Moreover, Frege goes on to suggest that because we can't be sure, successful communication would not be possible if it depended on mental representations being associated with words. Successful communication requires that we can rule out the possibility of mismatches. As Alexander Miller puts it, "there is nothing in our everyday conceptions of communication and grasp of sense that allows for such a possibility… the admission of such a possibility effectively renders the notion of sense empty" (Miller 1998, p. 40-1). But the problem with this argument is that, on the face of it, successful communication *doesn't* require that people can always establish that they are talking about the very same thing. What matters is simply that they *are* talking about the same thing, not that they *know* that they

are, much less are *certain* that they are. That is, what's important is that speakers and hearers are reliably coordinated in how they interpret one another's sentences. If it were a matter of mere chance that they interpreted a sentence in the same way, then perhaps we should say that they weren't communicating. But there is no reason to think that this is how things are. Of course, coordination may occasionally break down, and then communication deteriorates. But in the normal course of events, it doesn't break down and is maintained by mechanisms of language acquisition that contingently yet reliably attach mental representations with the same content to the same words. Why think that communication requires anything more?[11]

Frege's specific objections to mental representations are misplaced. However, contemporary proponents of the Semantic View aren't always motivated by concerns about communication. Peacocke (2005) has recently argued that concepts cannot be mental representations because this would rule out the possibility of there being concepts which human beings did not, or could not, ever entertain.

> If we accept that a thinker's possession of a concept must be realized by some subpersonal state involving a mental representation, why not say simply that the concept is the mental representation? Just this proposal is made by Margolis and Laurence (1999, p. 77). Mental representations that are concepts could even be typed by the corresponding possession condition of the sort I favour. This seems to me an entirely legitimate notion of a kind of mental representation; but it is not quite the notion of a concept. It can, for instance, be true that there are concepts human beings may never acquire, because of their intellectual limitations, or because the sun will expand to eradicate human life before humans reach a stage at which they can acquire these concepts. "There are concepts that will never be acquired" cannot mean or imply "There are mental representations which are not mental representations in anyone's mind". If concepts are

[11] These same considerations argue against Michael Dummett's case against concepts being mental representations, which he often discusses in the context of rejecting the "code theory of language" (viz., the view that words inherit their content from the mental representations with which they are associated). 'If this were the true account, communication would indeed rest on faith. Since it obviously cannot rest of faith, there must be an adequate outward manifestation of understanding…' (Dummett 1989, p. 187). The problem with this argument is that it *isn't* obvious that communication requires anything more than reliable correspondences between language and thought. If that's to say that communication rests of faith, then so be it. What's more, we'd suggest that one ought to be deeply suspicious of any theory of language that does what Dummett seems to want, viz., entirely preclude the possibility of miscommunication. Although it is very unlikely that people massively and systematically fail to communicate, it does seem (to us anyway) that miscommunication is a fact of life.

> individuated by their possession conditions, on the other hand, there is no problem about the existence of concepts that will never be acquired. They are simply concepts whose possession conditions will never be satisfied by any thinkers. (Peacocke, 2005, p. 169).

This argument embodies much the same type-token confusion as the one found in Frege. Peacocke's observation that there are concepts that will never be acquired is readily accommodated by the Psychological View. Recast in the framework of the Psychological View it simply amounts to the claim that there are mental representation types that will never be instantiated in anyone's mind.[12]

Earlier we saw another form of argument along these general lines, according to which the same mental representation can be associated with different concepts (Rey 1994), or the same concept with different mental representations (Peacocke 1992), and, consequently, that concepts can't be mental representations. Here we don't have a type-token confusion so much as a failure to fully appreciate that there are multiple ways of typing the very same tokens. For example, the argument that the same representation, say, an image of Paris, might correspond to thoughts of both Paris and France implicitly assumes that various image tokens are to be typed according to what they "look" like.[13] Then it's hard to see how there could be two thoughts—a Paris-thought and a France-thought—which look the same. But we could just as easily type an image not in terms of what it looks like but in terms of *what it is about*. Then the two image tokens would count as different concepts owing to their differing contents, just as two words can sound alike yet have different contents ("bank" the side of a river vs. "bank" the financial institution). The same considerations apply going in the other direction as well. Just as two words typed non-semantically can have the same content (the English word "cat" and the French word "chat"), different mental representations typed non-semantically can have the same content. Once again the argument against mental representations ignores the fact that there are different ways of typing mental representation tokens (Laurence & Margolis 1999).

---

[12] Like Frege, Peacocke seizes on the idea that every instance of a mental representation (i.e., every token) occurs in someone's mind. While this is true, there is nonetheless, no more reason to think that every mental representation type must be instantiated than there is to suppose that every property must be instantiated.

[13] For sake of argument, we are assuming that one and the same image can serve as a representation of Paris and a representation of France. But the mere fact that a single image is associated in introspection with two distinct thoughts shows very little. It might be that the thoughts are subserved by distinct non-imagistic representations and that the image is just vividly linked with each of the thoughts.

We take it that this insight—that there are different ways of typing mental representation tokens—opens the way to reconciling the two dominant theories of concepts. They can be combined by maintaining that concepts are mental representations *typed in terms of the senses they express*. We call this the Mixed View of concepts.

The Mixed View of Concepts

Concepts are mental representations typed in terms of the senses they express.

The Mixed View aims to combine the explanatory strengths of the two dominant traditions in the study of concepts. Following the Psychological View, it wholeheartedly embraces mental representations. Mental representations explain the productivity of thought and the fact that mental processes can be, at once, physical and rational. At the same time, following the Semantic View, the appeal to senses provides an explanation for the fact that two expressions (including two mental expressions) can have the same referent yet differ in cognitive significance. The suggestion is that mental representations can present a referent in different ways in virtue of expressing different senses. Senses continue to be intermediary entities, standing between expressions and referents, only now the expressions in question occur in an internal system of representation.[14]

Much of the rest of this paper will be devoted to evaluating the Mixed View of concepts. At this point, we take it that we have established that the Mixed View is a coherent option; we see no inherent obstacle to combining mental representations with senses. But that isn't to say that the Mixed View is preferable to either of the other views that it draws upon. In sections 4-7 we'll focus on challenges to the Mixed View that are associated with its commitment to mental representations. In section 8, we'll focus on challenges associated with its commitment to senses.

[14] The reason the Psychological View and the Semantic View can be combined in this way is that the Psychological View doesn't say that mental representations *are* meanings. Rather, it says that they *have* meanings, and leaves it as an open question what meanings are and what conditions have to be met to have a meaning. As a result, the Psychological View is perfectly consistent with the claim that meanings are Fregean senses. It's worth nothing, however, that that the Psychological View is also compatible with the claim that there are no such things as meanings (i.e., the view that while mental representations and sentences have meaning, there are no entities—meanings—needed to explain this fact.)

4. The Regress Argument

The first objection we will consider is an objection to any theory of concepts that incorporates mental representations. The objection is that mental representations are idle in that they reintroduce the very sorts of problems they are supposed to explain.

Simon Blackburn expresses the viewpoint well when he refers to any commitment to mental representations as a *dog-legged* theory (Blackburn 1984, p. 40). His complaint is that, when certain puzzling properties of natural language are purportedly explained by invoking mental representations, the mental representations that are invoked have exactly the same puzzling properties—so no progress is made. For instance, he asks how mental representations can explain the semantic properties of natural language if one then has to explain the semantic properties of mental representations themselves. Is one to introduce yet another level of representation to explain the semantic properties of the originally postulated mental representations? As he sees it, the theorist who appeals to mental representations is trapped in a dilemma. Either she is involved in an infinite regress (layer upon layer of representations), or else she should abandon mental representations altogether and directly explain what needs to be explained about language. Either way, mental representations seem to be a bad idea.

This form of argument traces back to Wittgenstein's later philosophy and has been enormously influential.[15] Nonetheless, the argument doesn't work, and the widespread mistrust of mental representations that it has engendered is unjustified.

One reason for rejecting the regress argument is that the problems, as they recur for the internal system of representation, may well take on a slightly different character or be open to a somewhat different range of potential solutions. An internal system lends itself to certain theoretical possibilities that are not at all promising for natural language. For example, nomic-causal theories of content say that the content of a symbol is a matter of what reliably causes it to

---

[15] For example, here's Michael Dummett (1978, p. 98; italics added):

> We know what it is to associate a word of one language with a word of another: asked to translate the one word, he utters, or writes down, the other. But the concept [i.e., the sense] has no representation intermediate between it and its verbal expression. Or, if it does, we still have the question what makes it a representation of that concept. We cannot say that someone's association of a particular concept with a given word consists in the fact that, when he hears that word, that concept comes into his mind, for there is really no sense to speaking of a concept's coming into someone's mind. All that we can think of is some image's coming to mind which we take as in some way representing the concept, *and this gets us no further forward, since we still have to ask in what his associating that concept with that image consists*.

occur. It's widely understood that such a theory is a nonstarter for natural language. Seeing a bird doesn't cause people to say "that's a bird". However, seeing a bird may cause people to *think* THAT'S A BIRD. So it is a substantive question whether the representation BIRD has its content in virtue of its causal link with birds. This is just one example, but it illustrates the general point. To the extent that the explanatory prospects are better for an internal system of representation, the regress argument fails. We may very well have good reason to posit an internal system of representation and stop there, with no comparable reason to posit further levels of representation. Then far from being idle, mental representations would turn out to be a key component in a full treatment of natural language (Fodor 1975, Crane 1995).

The second problem with the regress argument is more fundamental and tends to get overlooked (Laurence & Margolis 1997). One of the things that Blackburn and like-minded philosophers ignore is that there are independent reasons for introducing mental representations, reasons that aren't inherently tied up with the aim of explaining features of natural language. For example, we saw earlier that the productivity of thought bears explaining and that RTM does explain it by reference to the combinatorial structure of the internal system of representation. Likewise, RTM contributes to an explanation of an enormous number of details concerning the character of thought—patterns of error, breakdown, and a host of quantitative effects—that have emerged through the scientific study of the mind. But the whole force of the regress argument depends on there not being any independent grounds for positing mental representations. For once it is granted that mental representations are needed, there is nothing to stop a theorist from appealing to them in other explanatory projects. Philosophers who appeal to them in an account of language aren't dreaming them up with the sole purpose of delaying an explanation that ought to be given directly for natural language. Rather, they are taking advantage of an explanatory construct that is well justified separately from its role in theories of language. And given that the problems for language and thought are quite similar, it's hardly a stretch to suppose that the explanations for natural language go through explanations for mental representation.[16] Of course, whether such accounts should be pursued is a matter of how fruitful the explanations turn out to

[16] Though there is much philosophical argument about which is the more explanatorily basic domain between language and thought, a strong case can be made that thought is the more basic of the two. For one thing, there is an enormous body of evidence for thought processes in non-linguistic creatures—infants, animals, and severely aphasic adults (see, e.g., Baillargeon 2002, Clayton, Bussey, & Dickinson 2003, Varley 1998). For another, it's only under the assumption of thought's priority that we can explain how language can be learned (Fodor 1975, Pinker 1994).

be. But that's just to say that RTM-based approaches to explaining such things as the semantic properties of natural language should be assessed according to the standard measures in philosophy and science and that they aren't be dismissed out of hand. In spite of its apparent attraction, the regress argument poses a pseudo problem that raises no difficulties at all for theories of concepts, like the Mixed View, that incorporate mental representations.

5. Unstructured Mental Representations?

The second objection to the Mixed View is also directed toward its commitment to mental representations. This time the objection isn't that mental representations are idle. It's that the internal system of representation lacks the necessary structure for identifying concepts with mental representations. In particular, mental representations, according to this objection, don't form a combinatorial system with a compositional semantics.

Consider the thought that birds fly and the thought that birds eat. It's reasonable to say that these are similar in content insofar as they both employ the concept BIRD. According to the view behind the current objection, however, the mental representations corresponding to these thoughts don't share a constituent. They manage to pick out similar contents without containing a token of the same representational type. The problem, then, is that if concepts are mental representations, we can't say that these two thoughts involve a shared concept; by hypothesis, they don't. So the argument suggests that we are better off not identifying concepts with mental representations to begin with.

There are a number of views that take mental states to lack the relevant sort of structure, including some functionalist theories (e.g., Loar 1981), adverbial theories (e.g., Tye 1989), and some connectionist theories.[17] However, it is not at all clear that such views can be made to work. They face a host of difficult problems, not the least of which is the need to explain productivity. We saw earlier that RTM accounts for productivity by invoking *structured* mental representations. Without this structure, it remains deeply mysterious how a finite mind can be productive. In addition, unstructured representations lead to a huge burden in providing accounts

[17] It would be a mistake to lump together all connectionist theories as standing in opposition to versions of RTM that implicate structured representations (Aydede 1997, Marcus 2001). Nonetheless, connectionist theories are often adopted on the grounds that they have no need for structured representations.

of what appear to be structure-sensitive cognitive processes and cases of rule-based learning—psychological phenomena that are readily explained by models that incorporate structured representations (Marcus 2001, Gallistel & Gibbon 2002). This isn't to say that unstructured representations have no role to play in theories of the mind. But it's doubtful that a general account of thought can do without structured representations.

Where does this leave the Mixed View? Theories that are limited to unstructured representations have their work cut out for them, and until we can start to see how a theory of the mind can get by with just unstructured representations, the present objection has little force. Still, one might want to press the philosophical question of what would happen if somehow this explanatory burden were met. Assuming that psychology were to figure out a way to get by with just unstructured representations, what would the right thing be to say about concepts?

It seems to us that there are two ways to go. The first is to say that the situation envisioned is not so much one where concepts fail to be mental representations as one in which *there simply aren't any concepts*. The model in question may be a form of intentional realism insofar as it incorporates contentful mental states with causal powers. But it denies that there are any causally efficacious components of mental states. Theorists who endorse a model of this kind may continue to talk about thoughts sharing something—sharing a "concept"—but this is arguably just a manner of speaking. The concepts in question would lack the causal standing that thoughts are taken to have, and, for this reason, a full account of the mind and its intentional properties could dispense with them altogether. This first response is the one we are partial to. Nonetheless, there is a second response, which attaches less significance to the matter of causal efficacy, shifting the emphasis away from mental representations altogether. Once this is done, the natural outcome is to identify concepts with abstract objects, as the Semantic View says.

Given the availability of the second response, one might think it's best to hedge one's bets and simply adopt the Semantic View come what may. After all, the Semantic View can be held across a wide variety of empirical outcomes and is neutral about how productivity and related phenomena are to be explained. Doesn't that make it the preferable theory? No, it doesn't, and for two reasons. The first is that one has to be cautious about opting for a theory simply because it is more general and thereby able to cover more possibilities. The problem is that generality has to be balanced against explanatory power and that at some point the loss in explanatory power isn't worth the cost (for more on this, see the next section). The second is that an insistence on

generality cuts both ways. Notice that there are many competing semantic theories on offer, and not all of them take propositions to have constituents. For example, possible world semantics takes propositions to be sets of possible worlds or functions from worlds to truth values.[18] The point is that it could turn out that a form of possible world semantics is the best semantic theory. Then in anticipation of this possibility, a drive for generality would dictate that we hedge our bets against the Semantic View as well. Clearly, too much generality isn't necessarily a good thing.

The conclusion of this section is more tentative than the last, but we have argued that the envisioned situation (in which there are no suitably *structured* mental representations) does not necessarily show that concepts—if they exist—are other than what the Mixed View says they are. Arguably the best thing to say about this situation is that concepts are the components of propositional attitudes, just as the Mixed View says, but that in such a situation there wouldn't be any concepts. We should emphasize, however, that the envisioned situation is very much a hypothetical one and that there are powerful reasons to suppose it will stay that way.

6. The Martian Argument

The next objection to the Mixed View is also an objection to mental representations. In this case, it's not that mental representations are explanatorily idle or that they lack sub-propositional constituents in human minds. The objection this time is that *philosophical* theories of concepts shouldn't be tied to the contingencies of human psychology at all.

While there are a number of related motivations here, they all trace back to a similar picture of philosophy as standing apart from scientific investigation. For some philosophers, the operative assumption is simply that philosophy is properly understood as an a priori discipline and, as a result, that its theories are independent of empirical inquiry. For others, the concern has more to do with a quest for generality. Philosophical investigation is understood to function at the most inclusive level, needing to take into account not just how things are, or how things are

[18] Though possible world semanticists don't take propositions to have sense-like constituents, they do usually posit sub-propositional meanings and take propositions to be compositionally determined on the basis of these. The main motivation for this is a learnability or representability constraint (see, e.g., Davidson 1965). But if psychology is able to do without combinatorially structured representations, as we are assuming for the present objection, then it is hard to see what would motivate such constraints.

here, but also the full range of possible cases, including the minds of Martians, angels, and God. Either way, the sentiment is that the facts regarding mental representations are to be relegated to empirical psychology and that they aren't of *philosophical* interest.

One of the difficulties in addressing this objection is that, if a philosopher insists that, as a philosopher, she is interested in theories that achieve maximal generality, there isn't much one can say in reply. It's hardly worth arguing about personal preferences or what strictly speaking counts as philosophy. However, if the objection is to be understood as a substantive criticism of the Mixed View, it should be pointed out that there is a considerable cost when our theories are guided by such a preoccupation with generality, namely, a loss in explanatory power. The more we demand that our theories cover and the more we stretch them to include all the various esoteric possibilities, the thinner and more superficial they become.

It should also be evident that explanatory power often has the final say. Consider, for example, how the situation looks with biological kinds such as cells. Surely it would be perverse for a philosopher of biology to insist that theories tied to human biology are too parochial and that we need to aim for greater generality in characterizing these kinds, taking into account possible tissues found in Martians or other merely possible beings. A biological theory that aimed for that kind of generality wouldn't just be wildly unconstrained; it would explain few, if any, of the basic facts of cellular processes. After all, Martian "cells" might have no nuclei or DNA and might be subject to such radically different biochemical principles that books on Martian "cell biology" wouldn't overlap in any significant way with books on human cell biology.

Somewhat closer to the study of concepts, the same tradeoff between generality and explanatory power occurs in linguistics. Linguists typically distinguish between natural languages and other systems of communication, for example, between Spanish, German, and Hindi, on the one hand, and vervet monkey calls, on the other (see Cheney & Seyfarth 1990 for an analysis of vervet calls). Now one could try to maintain that every system of communication constitutes a language and that linguistic theories ought to be as inclusive as possible, treating "Vervetese" and German as instances of a general kind that is the primary object of investigation. But there are sound reasons for not doing things this way. Natural (human) languages have distinctive properties. They are productive, they aren't tied to the here and now, and they are able to express an extraordinary range of contents—to name just a few. What's

more, despite their enormous complexity, they develop effortlessly in all normal children within a biologically constrained critical period and in the face of significantly impoverished data. A theory of language that misses out on these features of human language just doesn't make enough contact with the very things that render it of theoretical interest. From an explanatory point of view, Spanish, German, and Hindi are all of a kind; Vervetese is only related to these in minor and explanatorily insignificant ways. Treating Vervetese on a par with human natural languages would make it impossible to draw any but the most superficial conclusions about "language", depriving the theoretical kinds of linguistics of nearly all of their explanatory power.[19]

We think that that the same considerations argue against seeking a maximally general theory of concepts. Do we really want a theory of concepts to be so general in its scope that it can cover Martians and angels? Angels aren't even material beings, so there is little reason to expect that they have much in common with us. And all we know about Martians is that Martian psychology differs in some unspecified way from human psychology. In the face of such obscure and unarticulated possibilities, why would concepts be any different than cells?[20]

7. Neo-Behaviorism

The next objection to the Mixed View is the last one that is directed to the commitment to mental representations. The objection is rooted in behavioristic approaches to mental content. Though few philosophers these days are openly committed to behaviorism,[21] it's still not uncommon to

---

[19] Of course, there may still be *some* commonalities between human and Martian psychology, just as there may be *some* commonalities between human and Martian biology. If the line of reasoning that we advocate in the text is correct, it would be a mistake to express these points of contact by saying that Martians have concepts. But that doesn't mean we can't express them in some other way. It just means that we need to introduce more insipid theoretical constructs, ones with few explanatory ambitions. Likewise for Vervetese and English. If it is a mistake to say these are both natural languages, we can still refer to them as, for example, "systems of communication". We just need to remember that the idea of a system of communication doesn't explain very much.

[20] It is also worth mentioning that the question of whether a priori approaches to philosophical problems are viable is a substantive issue that may partly turn on the nature of our concepts and conceptual abilities. To assume that philosophical theorizing about concepts must exclude empirical considerations begs the question against those who see in the theory of concepts itself grounds for being suspicious of aprioristic methods.

[21] A straightforward behaviorist theory of concepts claims that concepts are behavioral dispositions—for example, that the concept SPOON is the disposition to interact with spoons in a characteristic publicly observable manner. One major problem with this view is that, once we turn to concepts that move beyond concrete middle-sized objects (e.g.,

find philosophical views that have strong behaviorist leanings. The most sophisticated view along these lines is owing to Daniel Dennett and his *interperetivist* approach to mental content (Dennett 1978, 1987, 1991).

Interpretivists maintain that theories of belief and other propositional attitudes must be constrained by the limited epistemic vantage point of an interpreter. For Dennett, the interpreter has access only to the agent's overt behavior, and her job is to attribute beliefs that make sense of the behavior—and predict future behavior—under the assumption that the agent is rational. In making these attributions, the interpreter is "taking the intentional stance" towards the agent. Now many theorists would claim that there has to be some connection between overt behavior and content; at the very least, overt behavior provides evidence regarding what a person believes. But the point of interpretivism is that the link is supposed to be far tighter. Overt behavior is supposed to be *constitutive* of a belief, not merely good *evidence* for it. So, for example, in one of Dennett's definitive statements of the view, he claims that beliefs are constituted by patterns in observable behavior that can be seen by adopting the intentional stance (Dennett 1991).

One consequence of Dennett's interpretivism is that by identifying beliefs with behavioral patterns, the view is completely neutral about events that occur inside a person's head. Interpretivism is consistent with there being mental representations, but, for all it cares, you could open up a person's head and find nothing in there at all. Or you could open up different heads and find radically different inner workings. Nonetheless, they would have the same (or similar) beliefs so long as their overt behavior warrants the same (or similar) interpretations. Supposing that this is right, one might nonetheless want to say that having a belief requires having a concept, for example, that the belief that cats eat and the belief that cats sleep share a concept (viz., the concept CAT). But if beliefs aren't mental representations, concepts aren't going to be mental representations either. Interpretivism stands to undermine the Mixed View of concepts.

What's more, several of the motivations for interpretivism that Dennett cites are direct challenges to the identification of beliefs and concepts with mental representations. One of these

---

DEMOCRACY, NUCLEUS, INSTITUTION), it isn't at all plausible that there are distinctive behavioral dispositions that go with having a concept. More important perhaps is the fact psychology needs a more robust notion of mental causation than behaviorist principles allow. At best, behaviorism limits us to causal dispositions, yet even simple cognitive tasks tend to require extensive psychological *processes*—complex chains where one mental state causes another (Fodor 1968).

is that the beliefs we have are supposed to outstrip what we've entertained. As an illustration, Dennett remarks that "it should come as no surprise to any of you that zebras in the wild do not were overcoats, but I hazard that it *hadn't occurred* to any of you before just now" (Dennett 1977, p. 104; italics in original). Dennett's point is that, while we all have this belief, there is no reason to think that we all have a corresponding explicit mental representation stored somewhere in our brains. He is also doubtful that we can explain the belief as being represented in some looser sense, for example, as being entailed by a set of core beliefs that are explicitly represented. The reason is that empirical investigation could reveal that the explicit representations don't correspond to *any* of the beliefs we attribute to one another. "[I]f you were to sit down and write out a list of a thousand or so of your paradigmatic beliefs, all of them could turn out to be virtual" (Dennett 1981, p. 56).

A second motivation that Dennett cites for his interpretivism is that we are often unable to say whether two belief attributions pick out one and the same state—for example, whether the belief that 2 is less than 3 is the same as the belief that 3 is greater than 2 (Dennett 1981, p. 55). And related to this is a third objection, that we are generally happy to attribute beliefs without regard to whether the underlying states are identical (Dennett 1981, p. 54-5):

> It seems to me that the evidence is quite strong that our ordinary notion of belief has next to nothing of the concrete in it. Jacques shoots his uncle dead in Trafalgar Square and is apprehended on the spot by Sherlock; Tom reads about it in the Guardian and Boris learns of it in Pravda. Now Jacques, Sherlock, Tom, and Boris have had remarkably different experiences—to say nothing of their earlier biographies and future prospects—but there is one thing they share: they all believe that a Frenchman has committed murder in Trafalgar Square.

Dennett goes on to remark that if we were to assume, in each case, the presence of the same mental representation—"a similarly structured object in each head"—that this would be "a gratuitous bit of misplaced concreteness" (Dennett 1981, p. 55). The jibe emphasizes the contrast between Dennett's interpretivism and any view that takes beliefs to be mental representations. As Dennett sees it, beliefs can't be any more concrete than patterns of observable behavior; likewise, one might think, for concepts.

Let's start with Dennett's first argument for his interpretivist approach, the one that says that the beliefs we have outstrip what we've actively entertained. We don't see that this presents an especially difficult challenge for a theory of concepts based on RTM. RTM is a theory of *occurrent* states, ones that actively enter into mental processes. However, the sort of example that Dennett has in mind is clearly not an occurrent belief, so RTM doesn't apply.[22] Similarly, Dennett is concerned that paradigmatic beliefs, ones we would write down when asked what we believe, may not correspond to the representations that psychology would uncover. But what RTM says is that when a belief (paradigmatic or not) is implicated in a psychological process then there is a corresponding representation. Dennett is simply hedging his bets one way, RTM the other. In this context it is also important to note that, while commonsense packs a lot into the notion of a belief, RTM, as an explanatorily driven program, needn't agree with commonsense on every point. So the paradigms for commonsense may include what you'd write down when asked what you believe. But the paradigms for RTM concern *what actually occurs in your mind when thinking and acting*. One of the disturbing facts that cognitive science has uncovered is that the two don't always correspond (Nisbett & Wilson 1977, Wilson 2002). It doesn't follow that beliefs aren't representations. All that follows is that our belief attributions—including our self-attributions—don't always track the causally efficacious states that underlie behavior.

Dennett's second argument for interpretivism is that it makes sense to attribute the same belief to people despite their varying experiences and, presumably, their varying representations. This argument isn't much better than the first. Dennett is certainly right that belief attributions don't always reflect cognitive differences between us. But the natural explanation of this fact has less to do with beliefs themselves than with the practice of *attributing* beliefs. Saying what someone believes in daily life is a complex activity, one that is subject to pragmatic constraints and to limitations owing to our ignorance of people's minds. Sometimes when we say that someone believes that *p*, we intend to convey that she has a belief with that precise content. Other times, we may not be in a position to be so precise, or may have little interest in such precision. Much depends on the purpose of the attribution, as well as our understanding of the

[22] Many of Dennett's opponents follow Dennett in assuming that, in advance of considering the matter, people really do believe things like *zebras in the wild don't wear overcoats*. A standard view is that these are "tacit" beliefs that somehow trace back to the contents of occurrent mental states (e.g., Lycan 1988). The problem with this strategy is spelling out the conditions on tacit belief without inadvertently including things that, by all accounts, people don't believe. A better strategy, adopted by Pat Manfredi (1993), is to deny that people have (e.g.) the zebra belief prior to the subject being raised. Instead, they are simply disposed to form the belief upon considering the matter.

needs and interests of our conversational partners. For example, suppose that Alice sees that Bob and Cathy are arguing about a parking spot and reports the event to a friend later in the day. Alice and the friend may both know that Bob doesn't know Cathy by name. Still, given that Alice and her friend both know Bob and Cathy, it makes perfect sense for Alice to say "Bob believed that Cathy stole his parking spot". Reporting the belief in this way is an effective means to convey the information relevant to the conversation even though Bob didn't represent the situation to himself in quite these terms. Of course, if Bob *did* know Cathy by name and recognized her as the person who stole his parking spot, then this too would be an appropriate situation to say "Bob believed that Cathy stole his parking spot". Under these conditions, the attribution might better convey the nuances of Bob's belief state. But the fact that the attribution is appropriate in both this case and in the original doesn't show that beliefs aren't mental representations. It just shows that the practice of belief attribution is sensitive to a variety of factors.

What about the difference between the belief that 3 is greater than 2 and the belief that 2 is less than 3? For some purposes, it simply doesn't matter which of the beliefs we attribute. We aren't interested, in these cases, in the kind of precision that would tease them apart. But in other cases, the difference may matter and may reveal itself in both subtle and (sometimes) not so subtle ways. We are all familiar with the fact that sentences involving double negatives are harder to process than simple assertions despite the fact that they have essentially the same content. There are also less well-known ways in which the framing of an issue can have a striking psychological impact. For example, in one study U.S. doctors were asked about an operation and whether they would recommend it. In one condition, the operation was described as having a mortality rate of 66% within a period of five years; in the other, it was described as having a survival rate of 34% after a period of five years. The facts are the same in either case, but the recommendations weren't. Doctors told about the survival rate, as opposed to the mortality rate, were far more likely to recommend the operation (McNeil et al., 1982).

In sum, the motivations that Dennett cites for interpretivism don't provide sufficient motivation for abandoning the Mixed View of concepts. We'd argue as well that interpretivism is a deeply problematic view and that it doesn't offer a viable alternative to theories that identify thoughts and concepts with mental representations.

For starters, the theory is circular. It aims to explain the beliefs and desires of people's minds by reference to an interpreter. But what exactly is an interpreter? The answer, of course, is that an interpreter is an agent, a being with contentful mental states. So Dennett ends up explaining the states of one mind in terms of how it looks to another. Without the promise of an independent account of the interpreter, the theory simply helps itself to the very kinds of states that it is supposed to explain.[23]

Suppose, however, that we put the circularity objection to the side and simply grant Dennett the availability of an interpreter. We still face the question of who counts as an interpreter and what to do in cases of disagreements among potential interpreters. This question matters because we have every reason to expect that a single behavioral pattern is open to different interpretations. Maybe when John reaches for an apple only to put it back, what's going on is that he is hungry and remembers he is on a diet. On the other hand, maybe what's going on is that he thinks about throwing it at a rival but loses nerve. Or maybe he is testing his reflexes, or admiring the apple's color, or just fooling around, or … The point is that there are endlessly many possibilities here, all of which are consistent with a perspective that treats John as a rational being. Suppose, then, that we have two interpreters opting for two different explanations. Who is to say which is right? There is a tendency in Dennett's writings to suggest that both can be right, that mental content is subject to a certain amount of indeterminacy. But notice that the type of indeterminacy illustrated by our example does not involve similar or related contents, such as the difference between thinking about snow and thinking about sleet—or even thinking about rabbits and thinking about undetatched rabbit parts (Quine 1960). On the contrary, the varying contents that different interpreters may settle on may be completely unrelated, without any observable behavior to tease them apart.

What generates the indeterminacy is that individual beliefs and desires don't come with their own distinctive behavioral patterns and hence that the object of interpretation must be entire sets of propositional attitudes (as Dennett happily acknowledges). The problem is that vastly different propositional attitude sets are bound to be consistent with the same pattern of behavior so long as appropriate adjustments are made within an interpretation. For example, two people can do different things even if they have the same operative desire so long as they believe different

[23] Another problem for Dennett's view, urged by an anonymous referee, is that it is unclear how the patterns of overt behavior that Dennett takes to be constitutive of beliefs can explain any behavior at all.

things about how to get what they want. Similarly, two people can do the same thing even if they have different desires so long as they have appropriate compensating beliefs. It's no use trying to rule out such cases by maintaining that an interpretation violates an assumption of rationality. The different desires may reflect nothing more than a difference in taste, and even among rational beings, there is lots of room to disagree about the best way to achieve an end. Moreover, the constraint that we interpret people's behavior according to a high standard of rationality is itself unreasonable given that we know that people don't live up to that standard and that their behavior is interpretable all the same.

The upshot of these considerations is that the interpretivist position entails not just a little or even a moderate amount of indeterminacy. It entails a *massive* amount of indeterminacy. In fact, it's hard to see that it rules out the simultaneous possession of virtually *every* belief by every agent given that, for just about any belief and any behavior, there is a propositional attitude set that includes that belief and predicts that behavior. As far as we can see, the cost of Dennett's interpretivism is way too high for it to seriously undermine the identification of beliefs and concepts with mental representations.

8. Objections to Senses

Up to this point, all of the objections to the Mixed View have been objections to its commitment to mental representations. But the Mixed View also helps itself to senses and claims that concepts are mental representations that are individuated in terms of the senses they express. It's time we took a look at this other dimension of the theory.

One objection to senses focuses exclusively on their status as abstract objects. Nominalists maintain that abstract entities of all kinds are mysterious things and that senses ought to be rejected along with numbers and universals. This isn't the place to evaluate the merits of nominalism as a general metaphysics, so we'll put such concerns to the side. Still, it's worth noting that many contemporary philosophers who have nominalistic leanings nonetheless recognize the existence of some abstract entities. Even Quine, despite his love of desert landscapes, recognizes the existence of sets on the grounds that they are indispensable for doing mathematics and science. For Quine and others who have followed him, the objection to senses isn't merely that they are abstract entities but that, unlike sets, they don't earn their keep; they

aren't explanatorily useful (Quine 1960, Harman 1967). Whether Quine is right about this is disputable. Certainly, proponents of senses think they are enormously useful. Further, it could be argued that that if Quine doesn't see the need for senses this is partly because his own approach to the study of mind and language is excessively behavioristic.

The objections we want to focus on for the rest of this section address the explanatory functions associated with senses but without the burden of Quine's behaviorism. We've seen that for proponents of the Mixed View, senses are motivated in connection with the mode of presentation problem (just as with the Semantic View). The problem, to repeat, is that co-referential expression can have differing cognitive value. The terms "George Orwell" and "Eric Blair" pick out the same individual, but someone could know both terms without realizing that "George Orwell is Eric Blair" is true, even though it's obvious that "George Orwell is George Orwell" is true. Similarly, if thinking occurs in an internal system of representation—as the Mixed View maintains—then there could be two mental representations, GEORGE ORWELL and ERIC BLAIR, which have the same referent yet differ in cognitive value. On the sense-based explanation that the Mixed View adopts, what makes this possible is that these representations express different senses and that they are typed according to which sense they express. In short: different sense, different cognitive value.

How good is this explanation? Ironically, while senses are hypothesized largely to solve the mode of presentation problem, there are at least two reasons for doubting that they can, in fact, solve this problem. Both of these turn on the fact that senses are mind-independent entities that are supposed to stand apart from us, like numbers or Platonic forms.

First, the sense-based solution to the mode of presentation problem says that the reference of a word or internal representation is mediated by a sense that we grasp. But what exactly does grasping consist in? Clearly, *grasping* is a metaphor for a cognitive relation that needs to be explicated. The problem is that it is hard to see how this can be done in a way that is consistent with the view that senses are abstract objects. Notice that the relation can't be causal, since senses, as abstract particulars, are supposed to fall outside the realm of physical causes and effects. But if it's not causal, the nature of the relation remains utterly mysterious. The upshot of this dilemma is that the Fregean solution to the mode of presentation problem is based on a promise that is unlikely to be kept. Without the hope of an explanation of how senses can be grasped, senses themselves aren't going to help as intermediate entities.

Second, if that weren't bad enough, senses aren't sufficiently different from ordinary referents to solve the mode of presentation problem. Senses, like most referents, are external to our minds, and because of this it's hard to see why we shouldn't be able to stand in different cognitive/epistemic relations towards them as well (Fodor 1998). Just as we can have different modes of presentation for a number (the only even prime, the sum of one and one, Tim's favorite number, etc.), we ought to be able to have different modes of presentation *for a given sense*. Or if we can't, then there ought to be a reason why we can't. But as Fodor points out, there doesn't appear to be any reason why senses themselves don't generate the mode of presentation problem. "Frege needs something that can both present referents to thought and individuate thoughts; in effect, he needs a kind of MOP [mode of presentation] that is guaranteed to have only one handle. He can't, however, get one just by wanting it; he needs to explain how there could be such things" (Fodor 1998, p. 19).

In short, senses are poor candidates for modes of presentation. But this is just another way of saying that senses don't explain the very phenomenon for which they were introduced. This suggests that, while the Mixed View of concepts is a perfectly coherent view, we might be better off dropping it, and seeing whether we can get by with the Psychological View.

9. Mental Representations and Modes of Presentation

Let's turn our attention, then, to the question of how the Psychological View might attempt to solve the mode of presentation problem. The Psychological View has several promising options for dealing with this problem. What these share is the idea that modes of presentation are to be identified with properties of mental representations. By psychologizing modes of presentation, there is no longer a puzzle about how they are grasped or why they can't be grasped in different ways. This is for the simple reason that there is no longer a gap between minds and modes of representations. Modes of presentation are directly built into our minds and how they function.

To organize the discussion, we will focus on two general approaches that identify modes of presentations with properties of mental representations. However, there are variations within each approach, and as will become clear soon enough, the differences across the two approaches may not be as large as they first appear. One strategy identifies modes of presentations with the

properties that constitute a mental representation's narrow content. The other identifies modes of presentations with certain formal properties of mental representations.

Let's start with the narrow content approach. This treatment of modes of presentation is familiar from work on conceptual role semantics. The common theme of much of this work is that the content of an internal representation is determined by (or partly determined by) its causal role in various cognitive and perceptual processes. In his classic defense of this approach to content, Ned Block emphasizes that conceptual role semantics is especially suited to the needs of psychology (Block 1986). One of his primary reasons for this claim is that psychological explanation often cares more about *how* people are thinking about a situation than about which objects and events enter into the truth-conditions of their thoughts.

Block's central example concerns indexical concepts. He asks us to consider the difference between a belief that he might express with (2) versus one he might express with (3):

(2) I am in danger of being run over.
(3) Ned Block is in danger of being run over.

Block is impressed by the different effect these states might have on one's psychology:

> Believing (3) cannot be guaranteed to have the same life-saving effect on my behavior as believing (2), since I may not know that I am Ned Block (I may think I am Napoleon). So there is an important difference between (2) and (3) with respect to causation (and therefore causal explanation) of behavior. (Block 1986, p. 619, numbering changed from original)

Notice, in addition, that there might be someone else, say, Smith, who has a belief that she'd express with (2). In this case, Smith's belief would have different truth conditions than the belief that Block expresses with (2), but the states might have much the same effect, causing Smith and Block to act in similar ways.[24] The result of these considerations is that psychology may best be served by a notion of content that fails to line up with the truth conditions of the states involved. Sometimes states with the same truth conditions should be treated as having differing contents,

[24] The similarity of effect is strongest if we imagine that Smith is Block's doppleganger in a Twin Earth scenario, but Twin Earth isn't needed to make the general point.

and sometimes states with different truth conditions should be treated as having the same contents. For Block, this is exactly what the notion of narrow content gives us. Narrow content is distinct from, and more fine-grained than, truth-conditional content. Moreover, Block has a particular proposal about what narrow content is. A mental state's narrow content is a matter of its conceptual role; the narrow content of a concept is a matter of how it contributes to the conceptual roles of the thoughts in which figures.

The notion of narrow content has been enormously controversial in recent years. Some critics of the notion claim that narrow content isn't really a type of content at all, while others argue that psychological generalizations are best couched in terms of broad content. We don't propose to settle such issues here. Rather, we simply want to draw attention to the way that conceptual roles can do much of the work that modes of presentation are required to do. Clearly, co-referential concepts can have differing conceptual roles. Just as the concepts I and NED BLOCK can have a different impact on thought so can GEORGE ORWELL and ERIC BLAIR, or THE MORNING STAR and THE EVENING STAR. With the first pair, one facilitates inferences about novels and literature, while the other doesn't; in the second, one facilitates inferences about the morning, while the other facilitates inferences about the evening. We don't have to appeal to anything external to the mind to explain these facts. The facts are fully accounted for by the roles associated with each concept. Thus one might argue that modes of presentation just are conceptual roles: differences in cognitive value are differences in conceptual role.

Unfortunately, things aren't so simple. This is because there are cases where modes of presentation appear to be more fine-grained than conceptual roles. Consider the following example. Bob hears someone at a party talking about a famous writer called "Noam Chomsky" who has written extensively about politics. He forms a CHOMSKY_1 concept. A short time later, Bob overhears another group of people talking about a famous writer called "Noam Chomsky" who has written extensively about linguistics and philosophy. Bob assumes that these are different people and forms a CHOMSKY_2 concept. Since Bob is not particularly interested in linguistics or politics, he soon forgets about the details associated with his two concepts and in each case remembers only that Chomsky is a famous writer. As a result, the referent of his two concepts is the same (viz., Chomsky) and the conceptual role is the same (both license the inference that Chomsky is a writer in addition to any inferences that commonsense dictates, e.g., Chomsky is a human being, eats food, breathes, etc.). Nonetheless, there is no reason why Bob

couldn't think $Chomsky_1 \neq Chomsky_2$, while recognizing that it is a platitude that $Chomsky_1 = Chomsky_1$. So conceptual role won't suffice to handle the full scope of the mode of presentation problem.

A natural suggestion at this point is to supplement conceptual role with other properties of mental representations, in particular, the formal properties that allow the cognitive system to reidentify tokens of the same representation type. These properties are akin to the orthographic properties of an ideal formal language; for this reason, we'll refer to them using the shorthand *mental orthography*. Advocates of conceptual role semantics can simply incorporate mental orthography into their account. One way is to claim that mental orthography is a determinate of narrow content along with conceptual role (i.e., modes of presentation = narrow content = [conceptual role + mental orthography]). Alternatively, another way is to claim that mental orthography is an aspect of modes of presentation that is independent of narrow content (i.e., modes of presentation = [narrow content + mental orthography]). Common to both of these proposals is the idea that modes of presentation amount to a combination of conceptual role and mental orthography.

We now come to the second general approach to psychologizing modes of presentation. This approach is motivated by considerations that are supposed to argue against any form of conceptual role semantics. The suggestion is that we identify modes of presentation exclusively with mental orthography and that conceptual role drops out of the picture altogether. Fodor (1998) adopts a theory along these lines. His primary argument against the conceptual role approach is that he thinks conceptual role semantics undermines the prospects of intentional psychology.[25] The problem is supposed to be that conceptual role semantics is excessively holistic since there is no principled way of distinguishing the inferences that count towards content from those that don't. The only principled view is to maintain that all of a concept's inferential relations figure into its content. But since people clearly don't share concepts with the same total conceptual roles, it follows that they can't literally share any concepts. Similarly, because people's total conceptual roles are constantly changing, they can't retain the same concepts over time. For Fodor, these forms of instability undercut the possibility of there being any true psychological generalizations that are defined over shared concepts. Fodor's theory of

[25] Notice that Fodor's motivation is directly at odds with Block's. Where Block thinks conceptual role semantics is needed for the purposes of psychological explanation, Fodor thinks that conceptual role semantics renders psychological explanation impossible.

concepts aims to avoid this consequence by not letting in any conceptual role whatsoever. For Fodor, content isn't determined by conceptual role; instead, it is determined by mind-world causal relations. Of course, we still need some way of teasing apart co-referential concepts. Fodor's solution is to say that modes of presentation aren't connected with content. A concept's mode of presentation is just its mental orthography.

While Fodor's sparse account of modes of presentation is intriguing, it faces a serious objection. Murat Aydede (1998) has argued that mental orthography is too weak to serve the function associated with modes of presentation, since mental orthography can't deliver conditions of interpersonal individuation. Consider the Chomsky example again. Bob's co-referential concepts, CHOMSKY_1 and CHOMSKY_2, are kept distinct via differences in mental orthography. But now suppose that Anne has a concept CHOMSKY_3 that also refers to Chomsky and shares the minimal conceptual role associated with both of Bob's concepts. (Like Bob, Anne heard that Chomsky is a linguist who has written many books, etc., but ends up only remembering that Chomsky is a famous writer.) Should we identify Anne's CHOMSKY_3 with Bob's CHOMSKY_1 or his CHOMSKY_2? There doesn't seem to be any way to settle the matter. The problem is that the causal properties that constitute mental orthography are almost certainly arbitrary within a broad class (e.g., the class of all names). As a result, there can't be any fact about whether Anne's representation is of the same orthographic type as one or the other of Bob's co-referential concepts.

How serious is this problem? It is worth noting that the amount of indeterminacy involved is quite limited. For example, it's nothing like the everyone-believes-everything sort of indeterminacy that plagues Dennett's approach to concepts and propositional attitudes. Moreover, indeterminacy in itself is not necessarily objectionable. What one should want from a theory of concepts isn't the complete eradication of all indeterminacy but rather that our theories match the extent to which our concepts are, in fact, determinate. We'd suggest that it really isn't so obvious which of Bob's concepts corresponds with Anne's. A theory that doesn't settle the matter may be exactly what the case demands.

Unfortunately, Aydede's objection is not limited to cases where people have identical conceptual roles associated with their concepts. Mental orthography will be equally indeterminate interpersonally when the associated conceptual roles differ. For instance, suppose

Anne's CHOMSKY$_3$ has conceptual role R1, while Bob's CHOMSKY$_1$ also has R1 and his CHOMSKY$_2$ has a very different conceptual role R2 (see figure 1).

Insert figure 1 about here

By hypothesis all three concepts have the same referent and there is no means for making interpersonal comparisons of mental orthography. So if Fodor only allows himself these two parameters to work with—reference and orthography—then there is no more basis for saying that Bob's CHOMSKY$_1$ is type identical with Anne's CHOMSKY$_3$ than saying that Bob's CHOMSKY$_2$ is type identical with Anne's CHOMSKY$_3$. The problem is that the resulting indeterminacy is no longer so easy to dismiss. The temptation is to say that the differing conceptual roles ought to break the tie and that Bob's two concepts aren't on equal footing when compared with Anne's. Yet Fodor's theory doesn't allow us to say this. Things get even worse if we consider a case where the conceptual roles implicate associated public language names. For example, suppose now that Anne's concept with role R1 is associated with the same name as one of Bob's concepts but that Bob also has a coreferential concept with R2 that is associated with a different name. It helps if we switch the example to an author with a pseudonym. So image that for both Anne and Bob the concept with R1 is associated with "George Orwell" and that Bob's coreferential concept with R2 is associated with "Eric Blair" (see figure 2).

Insert figure 2 about here

Once again, Fodor only has orthography and reference to work with. He can't say that one of Bob's concepts—the one associated with "George Orwell"—is the same as Anne's because of the significant correspondence in conceptual role. According to Fodor's theory of modes of presentations, the concept of Bob's that is associated with "Eric Blair" has an equally good claim

to being the same concept as Anne's. The result is that while mental orthography alone may distinguish concepts intrapersonally, it seems insufficient for dealing with interpersonal cases.

We believe that the best way for a theorist like Fodor to respond to these difficulties is to recognize a place for conceptual role without taking it to be constitutive of concepts. Even if we say that, strictly speaking, there is no determinate fact as to which of Bob's co-referential concepts is identical to Anne's, we can still explain the interpersonal similarities and differences that these concepts engender. Anne and Bob each have a concept with role R1. This role ensures that their concepts have much the same impact on their respective minds. Whether we go on to claim that R1 is a condition on the identity of their concepts doesn't matter. In other words, just because one denies that conceptual role is constitutive of a concept's identity doesn't mean one has to deny that concepts *have* conceptual roles. So Fodor can adopt the very explanatory strategy that Block uses but claim a difference in bookkeeping. Where Block would say that conceptual role is constitutive of a concept's narrow content, Fodor can say that the only content is referential content but also go on to say that the conceptual role associated with a concept is responsible for—and hence explains—the impact of that concept on how people think and act.[26]

Moreover, interpersonal indeterminacy needn't create confusion as people pass along information. Shared links to a public language will often ensure that the information is organized in similar ways in different minds. In the "George Orwell"/"Eric Blair" case, people generally have two concepts, one corresponding to each name. So when they hear something about Orwell under the name "George Orwell", they regularly place this information in connection with the concept they associate with "George Orwell", not the concept they associate with "Eric Blair". Strictly speaking, it may be indeterminate whether in doing so people are using pairs of concepts with the same modes of presentations. But in practice this makes no difference, since their conceptual systems are coordinated in that their respective pairs of concepts embody corresponding modifications to their patterns of inference.

In the end, it doesn't matter whether the machinery that is called on to explain mode of presentation effects is taken to be constitutive of concepts or not. All that matters is that concepts *have* the features that explain the effects. This insight has a variety of interesting consequences. One is that we should reconsider Fodor's argument that conceptual role semantics undermines

[26] Much the same might be said regarding the related problem of empty concepts. Conceptual role can also explain an empty concept's cognitive value even if the conceptual role is not taken to be concept constitutive—indeed, the explanation will be exactly the same as if it were constitutive.

psychology because of its holistic consequences. Fodor's concern is that holistic theories of content introduce instability in that they don't allow for different people to have the same concept or even for a single person to have the same concept at different times. While it remains true that differences in conceptual role introduce a form of instability, we can now see that the instability needn't be pernicious and that, in this respect, the Fodorian approach to concepts has little advantage over the conceptual role approach that it is meant to oppose. If we suppose that both approaches can provide equally stable accounts of reference, then the two approaches introduce the same kinds of stability and instability and merely label these effects in different ways.[27] For proponents of holistic forms of conceptual role semantics, changes in conceptual role amount to changes in narrow content, which in turn amount to changes in modes of presentation, and *these* explain the mode of presentation effects. For the revised version of Fodor's account, changes in conceptual role amount to changes in the inferences associated with concepts, and these explain the mode of presentation effects *directly* without going through narrow content or modes of presentation. Either way, however, the impact on a person's mind and interpersonal comparisons are explained in terms of the very same resources—reference, orthography, and conceptual role. How these are divvied up when we say what concepts are is little more than a bookkeeping issue.

The mode of presentation problem raises a number of complex issues that we don't have the space to pursue further, but for present purposes the main point is that there are more than enough options for a theory of modes of presentation that treats them as psychological properties of mental representations. The Psychological View can call upon conceptual roles and mental orthography and use these in various combinations. As a result, the mode of presentation problem isn't an obstacle to the Psychological View of concepts.

[27] This is not to say that the two accounts are entirely equivalent. They may differ on a number of further dimensions, including referential stability. Sometimes conceptual role theorists claim that they are happy to adopt the same mechanisms of reference determination as causal theorists like Fodor, but sometimes their theory of reference determination is supposed to have a different basis (e.g., one more akin to a description theory of reference). On the second type of strategy, the account of reference fixing may well introduce a greater degree of referential instability.

10. Conclusion

The guiding question of this paper is whether concepts should be identified with mental representations or abstract objects. After reviewing the principal motivations for these very different views, we showed how the two can be combined to accommodate a variety of related explanatory goals. The Mixed View speaks to the need for mental representations in explaining productivity, the nature of mental processes, and other basic facts about human minds. It also speaks to the need for modes of presentation. We saw, however, that senses are surprisingly poor candidates for doing the work that is expected of modes of presentation and noted several ways in which modes of presentation can be construed as psychological properties of mental representations. In light of the these options, the Psychological View can offer all of the explanatory benefits that the Mixed View promises but without the problems raised by taking modes of presentation to be abstract objects. So concepts should be identified with mental representations.

References

Aydede, Murat. (1998) "Fodor on Concepts and Frege Puzzles," *Pacific Philosophical Quarterly*, 79, 289-294.

Baillargeon, Renée. (2002) "The Acquisition of Physical Knowledge in Infancy: A Summary in Eight Lessons," in U. Goswami (ed.) *Handbook of Childhood Cognitive Development* (pp. 47–83), Oxford: Blackwell.

Blackburn, Simon. (1984) *Spreading the Word*, New York: Oxford University Press.

Block, Ned. (1986) "Advertisement for a Semantics for Psychology," in P. A. French, T. Uehling Jr., & H. Wettstein (eds.) *Midwest Studies in Philosophy, Volume X: Studies in the Philosophy of Mind* (pp. 615-678), Minneapolis: University of Minnesota Press.

Cheney, Dorothy, & Seyfarth, Robert. (1990) *How Monkeys See the World: Inside the Mind of Another Species*, Chicago: University of Chicago Press.

Clayton, Nicola, Bussey, Timothy, & Dickinson, Anthony. (2003) "Can Animals Recall the Past and Plan for the Future?" *Nature Reviews: Neuroscience*, 4, 685-91.

Crane, Tim. (1995) *The Mechanical Mind*, New York: Penguin.

Davidson, Donald. (1965) “Theories of Meaning and Learnable Language,” reprinted in D. Davidson (1984) *Inquiries into Truth and Interpretation* (pp. 3-15), New York: Oxford University Press.

Davis, Wayne. (2003) *Meaning, Expression, and Thought*, Cambridge University Press.

Davis, Wayne. (2005) *Nondescriptive Meaning and Reference*, Oxford University Press.

Dennett, Daniel. (1977) “A Cure for the Common Code,” reprinted in D. Dennett (1978) *Brainstorms* (pp. 90-108), Cambridge, MA: MIT Press.

Dennett, Daniel. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology,* Cambridge, MA: MIT Press.

Dennett, Daniel. (1981) “Three Kinds of Intentional Psychology,” reprinted in D. Dennett (1987) *The Intentional Stance* (pp. 43-68), Cambridge, MA: MIT Press.

Dennett, Daniel. (1987) *The Intentional Stance*, Cambridge, MA: MIT Press.

Dennett, Daniel. (1991) “Real Patterns,” *Journal of Philosophy*, 88, 27-51.

Dummett, Michael. (1978) “What Do I Know When I Know a Language?” reprinted in M. Dummett (1993) *Seas of Language* (pp. 94-105), New York: Oxford University Press.

Dummett, Michael. (1989) “Language and Communication,” reprinted in M. Dummett (1993) *Seas of Language* (pp. 166-87), New York: Oxford.

Fodor, Jerry. (1968) *Psychological Explanation*, New York: Random House.

Fodor, Jerry. (1975) *The Language of Thought*, Cambridge, MA: Harvard University Press.

Fodor, Jerry. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge, MA: MIT Press.

Fodor, Jerry. (1998) *Concepts: Where Cognitive Science Went Wrong*, New York: Oxford University Press.

Frege, Gottlob. (1892a) “On Concept and Object,” in P. Geach and M. Black (eds.) *Translations form the Philosophical Writings of Gottlob Frege*, 3rd Edition (pp. 42-55), Oxford: Blackwell.

Frege, Gottlob. (1892b) “On Sense and Reference,” in P. Geach and M. Black (eds.) *Translations form the Philosophical Writings of Gottlob Frege*, 3rd Edition (pp. 56-78), Oxford: Blackwell.

Gallistel, Charles R. & Gibbon, John. (2002) *The Symbolic Foundations of Conditioned Behavior,* Matwah, NJ: Lawrence Erlbaum Associates Inc.

Harman, Gilbert. (1967) "Quine on Meaning and Existence, I: The Death of Meaning," *Review of Metaphysics*, 21, 124-51.

Jackson, Frank. (1998) *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, Oxford: Oxford University Press.

Laurence, Stephen & Margolis, Eric. (1997) "Regress Arguments Against the Language of Thought," *Analysis*, 57, 60-66.

Laurence, Stephen & Margolis, Eric. (1999) "Concepts and Cognitive Science," in E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core Readings* (pp. 3-81), Cambrdige, MA: MIT Press.

Laurence, Stephen & Margolis, Eric. (2003) "Concepts and Conceptual Analysis," Philosophy & Phenomenological Research, 67.2, 253-282.

Loar, Brian. (1981) *Mind and Meaning*, Cambridge: Cambridge University Press.

Lycan, William. (1988) *Judgement and Justification,* Cambridge: Cambridge University Press.

Manfredi, Pat. (1993) "Tacit Beliefs and Other Doxastic Attitudes," *Philosophia*, 22, 95-117.

Marcus, Gary. (2001) *The Algebraic Mind: Integrating Connectionism and Cognitive Science*, Cambridge, MA: MIT Press.

McNeil, Barbara, Pauker, Stephen, Sox, Harold, & Tversky, Amos. (1982) "On the Elicitation of Preferences for Alternative Therapies," *New England Journal of Medicine*, 306, 1259-1262.

Miller, George. (1995) Interview in *The Human Language Series*, Equinox Films, directed by Gene Searchinger.

Miller, Alexander. (1998) *Philosophy of Language*, London: UCL Press.

Nisbett, Richard & Wilson, Timothy. (1977) "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review*, 84.3, 231-59.

Peacocke, Christopher. (1992) *A Study of Concepts*, Cambridge, MA: MIT Press.

Peacocke, Christopher. (2005) "Rationale and Maxims in the Study of Concepts," *Noûs*, 39.1, 167-178.

Pinker, Steven. (1994) *The Language Instinct,* London: Penguin.

Quine, W.V.O. (1960) *Word and Object*, Cambridge, MA: MIT Press.

Rey, Georges. (1994) "Concepts," in S. Guttenplan (ed.) *A Companion to the Philosophy of Mind* (pp. 185-193), Oxford: Blackwell.

Sutton, Jonathan. (2004) "Are Concepts Mental Representations or Abstracta?" *Philosophy and Phenomenological Research*, 68.1, 89-108.

Tye, Michael. (1989) *The Metaphysics of Mind*, New York: Cambridge University Press.

Varley, Rosemary. (1998) "Aphasic Language, Aphasic Thought: An Investigation of Propositional Thinking in an Apropositional Aphasic," in P. Carruthers & J. Boucher (eds.) *Language and Thought: Interdisciplinary Themes* (pp. 128-45), Cambridge: Cambridge University Press.

Wilson, Timothy. (2002) *Strangers to Ourselves: Discovering the Adaptive Unconscious*, Cambridge, MA: Harvard.

Zalta, Edward. (2001) "Fregean Senses, Modes of Presentation, and Concepts," *Philosophical Perspectives, 15, Metaphysics*, 335-59.

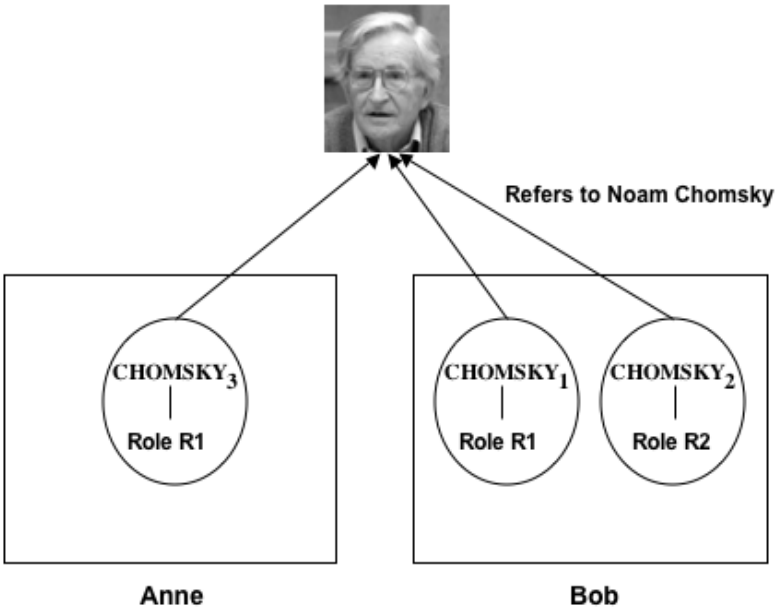

Figure 1. Fodor's theory can't say which of Bob's two concepts is identical with Anne's. For Fodor, concepts are individuated solely in term of reference and mental orthography (conceptual role is irrelevant), but mental orthography doesn't allow for interpersonal comparisons. (Photo © LSE/Nigel Stead, used with permission).

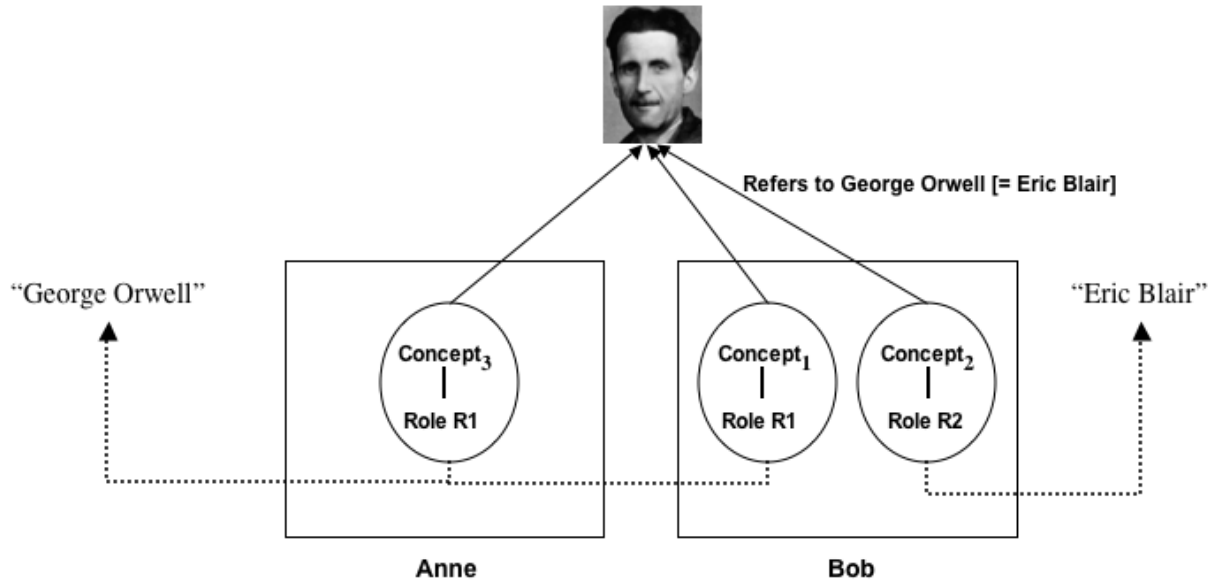

Figure 2. While Bob's two concepts have the same referent as Anne's, only one shares the same conceptual role and the same connection to a public language name. Nonetheless, Fodor's theory of concepts can't distinguish among Bob's concepts by appealing to these similarities. So, once again, Fodor's theory can't say which of Bob's two concepts is identical to Anne's. (Photo in public domain under terms of GNU General Public License).