Demonstrative Thought

JOSEPH LEVINE

Abstract: In this paper I propose a model of demonstrative thought. I distinguish token-demonstratives, that pick out individuals, from type-demonstratives, that pick out kinds, or properties, and provide a similar treatment for both. I argue that it follows from my model of demonstrative thought, as well as from independent considerations, that demonstration, as a mental act, operates directly on mental representations, not external objects. That is, though the relation between a demonstrative and the object or property demonstrated is semantically direct, the mechanism by which a demonstrative acquires its referent involves mediation by a perceptual representation. Finally, I argue that so-called 'demonstrative concepts'—which I treat as type-demonstratives—cannot perform the various philosophical functions that have been assigned to them.

1. Introduction

Suppose a fly is buzzing around the room. I've noticed it and it's bothering me. I suddenly see it on the wall and I say to myself, 'That fly sure is annoying!'. I've just expressed a thought, and a constituent of that thought is expressed by the phrase 'that fly'. The thought in question I'll call a 'demonstrative thought', and the thought-constituent expressed by the demonstrative phrase 'that fly' I'll call a mental demonstrative.

My goal in this paper is to provide an account of mental demonstratives. This goal is motivated by two principal concerns. First, for reasons which will become clearer as we go on, but might already be obvious to some, mental demonstratives are a plausible point at which to locate the mind-world interface, the point at which thought most directly touches the world. So any theory of intentionality, of how it is that thoughts are about what they are about, is going to pay special attention to the phenomenon of mental demonstration. In particular, I want to investigate to what extent, and in what sense, mental demonstratives 'directly', or 'immediately' pick out objects. Another way to put it is this. An intuitive idea is that sometimes we are directly 'acquainted' with an object—we just directly have it in mind, knowing which one it is without having to know some identifying

Address for correspondence: Department of Philosophy, 352 Bartlett Hall, University of Massachusetts, 130 Hicks Way, Amherst, MA 01003-9269, USA
Email: jle@philos.umass.edu

description—and demonstratively picking it out seems to be the mechanism by which we are afforded this sort of access to an object. I want to see whether, and to what extent, this intuition can be sustained.

Second, the term 'demonstrative concept' has crept into the literature in recent years, and the notion has been put to work to solve various philosophical problems. I have always been suspicious of the notion, unclear what the term really means, and doubtful that, whatever it is, it can do the work it's been recruited to do. But before one can determine whether these suspicions are warranted, it's necessary to get some account of demonstrative thought. Hence both my concern for understanding the mind-world interface in general and my suspicions about demonstrative concepts converged on the present project.

My plan for the paper is as follows. In section 2 I will set out a series of distinctions that will play a crucial role in my account of demonstrative thought. In section 3 I will present the account itself. Finally, in section 4 I will exploit my account to criticize three specific uses to which the notion of a demonstrative concept has been put in recent literature.

Before turning to the series of distinctions in section 2, it's important that I be explicit at the outset about an underlying commitment that constrains the rest of the paper: I presume what Jerry Fodor calls the 'Representational Theory of Mind' (RTM).[1] That is, intentional mental states are functional, or computational relations to mental representations, and their intentional properties are inherited from the intentional properties of the relevant representations. Mental representations on this view are conceived as structured mental particulars; they are often characterized as involving a 'language of thought'. So mental demonstratives will in the first instance be understood to be mental particulars, and the task will be to provide an account both of how they acquire their content and of how they play their special cognitive role. I will not provide a defense of RTM itself, except to say that, as Fodor and others have argued, it provides the best model of how thought could be materially realized and also seems to be the working assumption of much of cognitive psychology. I hope even for those who don't buy RTM it is sufficiently interesting to see what resources it has for explaining demonstrative thought and constraining notions like 'demonstrative concept' and 'directly thinking of an object'.

2. Some Distinctions

2.1 Linguistic versus Mental Demonstratives
I have already touched on this distinction, but I want to just briefly reiterate it. Utterances of 'That fly is sure annoying' contain demonstrative expressions, and their semantic behavior is much discussed in the philosophy of language. I will be drawing on some of that literature here (particularly Kaplan, 1989a,b), but it's still

---

[1] The *locus classicus* for his account is Fodor, 1975.

crucial to keep in mind that my topic is the kind of thought-constituent expressed by this linguistic expression, not the expression itself. That there are thought-constituents of this sort—what I'm calling mental demonstratives—follows fairly naturally from RTM. But though it's natural to assume, as part of RTM, that the semantic features of the linguistic expressions of our thoughts derive from the semantic features of the mental representations underlying them, this doesn't entail that the two can't ever come apart. Public conventions and other pragmatic or contextual features could result in the semantic content of an utterance differing from the semantic content of the mental representation that produced it. I am agnostic on the question whether this ever happens. I just want to emphasize that it's *mental* demonstratives, thought-constituents, that concern me here. Thus whenever I use the term 'demonstrative' in what follows, unless context or explicit qualification make clear otherwise, it's mental demonstratives I intend.

2.2 Token- versus Type-Demonstratives

A token-demonstrative picks out an individual, a token of some type, though type information needn't be part of the content of the expression.[2] 'That fly' obviously refers to a token of the type, or kind, fly. But I could also refer to it just as 'that', without even knowing it's a fly. For instance, looking at something on the wall I can ask, 'what is that?' I needn't have any particular sortal in mind in order to demonstrate it.

A type-demonstrative is a demonstrative expression that picks out a type or kind, as in 'that color' said while holding up and visually attending to a paint chip. One isn't referring to the actual paint on the surface of the chip, but to the color type of which that surface is a token. Now, just as I assume there is a constituent of thought, a mental representation, that corresponds to the token-demonstrative 'that (fly)', let's assume there is also a thought constituent corresponding to the type-demonstrative 'that color'. As I'll make clear in section 4, it is type-demonstratives that philosophers (usually) have in mind when speaking of demonstrative concepts.

Within the class of type-demonstratives, I want to distinguish between what I'll call 'true type-demonstratives' and 'pseudo type-demonstratives'. The distinction is easily seen by noting an intuitive difference between two different ways of demonstrating a type, a difference that McDowell (1996) emphasizes in his discussion of demonstrative concepts, to which I will return below. On the one hand, I can refer to the color of a house by looking at it and thinking 'that color', the sort of case we've been discussing. On the other hand, I can refer to the color of the house by looking at the house in the dark, not seeing its color, and still thinking 'that color',

---

[2] The terminology can be confusing since it can seem as if 'token' modifies 'demonstrative' in the sense that we're talking about a token of that demonstrative, like an utterance. Rather, what's meant is a demonstrative whose content is a token, or individual. Since the term 'type-demonstrative' is already in the literature (Loar, 1997), I decided to stay with this usage.

or, perhaps more precisely, 'the color of that house'. In both cases I've referred to the color of the house via a perceptual interaction with the house. Yet, while those who, with McDowell, talk about demonstrative concepts of properties or kinds would sanction the first way as a legitimate demonstrative concept of the color, they wouldn't classify the second one that way. What precisely is the difference?

Of course the difference is already revealed in the fact that, whereas the first way of referring to the color is quite naturally expressed with the type-demonstrative expression 'that color', the second one, as just mentioned, is really better captured with the more complex 'the color of that house'. What this reveals is that in the second case the demonstrative reference is really to the house—a token-demonstrative—and the color is referred to by way of a straightforward definite description incorporating the token-demonstrative. Pseudo type-demonstratives, then, are really disguised descriptions containing token-demonstratives, whereas true type-demonstratives involve a genuinely different sort of mental demonstration. From now on, when speaking of type-demonstratives, it is the true ones I intend.

2.3 Semantic versus Meta-semantic Questions

Theories of intentionality need to address two kinds of question, what Kaplan (1989b) has called the 'semantic question' and the 'meta-semantic question.' The former concerns the specification of the content of a representation; what are its truth conditions? how does it behave in modal contexts? The latter concerns the conditions by virtue of which the representation has the content it has; what makes it the case that this representation has this content?

For example, on a Kripkean account of proper names (Kripke, 1980), the semantic question is answered by the direct reference theory. The name 'Socrates' functions to refer to the guy, Socrates, and the referent is its semantic value. However, there is a longer account, involving dubbing ceremonies and historical transmissions, which answers the meta-semantic question; namely, how it is that our use of 'Socrates' now refers to the ancient Greek philosopher. Similarly, on a nomic-covariational theory of conceptual content, such as Fodor's (1990) or Dretske's (1981), one can distinguish between the answers to the semantic and meta-semantic questions. That is, on such a theory the semantic value of my mental representation [water][3] is the property of being water, which is the same as the property of being H_2O. In other words, they have a direct reference theory for concepts (or many concepts, at any rate).[4] The part of the theory that deals with nomiccovariation, however, is to answer the meta-semantic question. So, the story goes, [water] in my head refers to the property of being water because there exists the appropriate nomic connection between instantiations of the property of being water and tokenings in my head of [water].

---

[3] I will use brackets, instead of quotes, to mention mental representations, as opposed to terms of natural language.

[4] For Fodor, just about any concept that is expressed with a lexically simple term in natural language is treated as directly referential.

Notice that it's possible for representations to be accorded the same semantic treatment even though they are given different meta-semantic treatments. For instance, Kripke allows that names can be introduced by reference-fixing descriptions, though he thinks this isn't in fact the way most proper names in natural language acquire their reference. I can, to use Evans's (1979) example, name whoever is the inventor of the zip 'Julius'. While 'Socrates' refers to Socrates by virtue of the historical transmission leading back to an initial baptism, 'Julius' refers to whoever invented the zip by virtue of his/her satisfying the relevant description. These are two very different meta-semantic mechanisms. However, one can still maintain, as Evans does, that 'Julius', like 'Socrates', is directly referential, taking the person who invented the zip as its semantic value, and not the description by which the reference was fixed. Thus 'Possibly Julius didn't invent the zip' is true on this theory. The semantic story is the same for the two names, though the meta-semantic stories differ in crucial ways.

For the purposes of this paper, I am going to just assume a directly referential account of the semantics of demonstratives. The content of [that fly is annoying] is a singular proposition with the fly itself and the property of being annoying as constituents. Also, [that fly is currently in my visual field] is contingent, as the fly could have been elsewhere, though I would not have been able to demonstrate it if it had been. However, as the cases of 'Socrates' and 'Julius' demonstrate, having a directly referential semantics is consistent with a range of meta-semantic mechanisms. My concern in this paper is to understand the meta-semantics of demonstrative thoughts; in particular, perceptually based demonstrative thoughts, as exemplified by [that fly is annoying].

2.4 Intentionally Mediated versus Direct Meta–Semantic Mechanisms (IMM versus DMM)

By an IMM I intend a mechanism for securing reference that depends essentially on the content of another representation. So, for instance, even though 'Julius' is a directly referential term, with no descriptive material as part of its content, it nevertheless gains its reference through a description, and therefore counts as possessing an IMM. It might be thought that the relevant contrast here is 'Socrates', at least on Kripke's theory, since he argues that most proper names do not refer by virtue of a reference-fixing description. However, even though the meta-semantic mechanism for 'Socrates' is quite different from that for 'Julius', as we saw above, I would claim that it still counts as an IMM. After all, if we follow Kripke's story about historical transmission, it's clear that our current use of 'Socrates' acquires its content from our intention to use the term in conformity with those from whom we acquired the name. Thus there is a lot of intentional material involved in endowing the name with content. Notice especially how, on the Kripke story, the referent of 'Socrates' as used by us depends crucially on the initial baptism. Well, how does that work? Presumably, someone, the parent, says, or thinks, I will call this baby 'Socrates'. So the reference of 'Socrates' derives from the reference of 'this

baby', and therefore the meta-semantic mechanism fits the condition for being an IMM.

To be sure, Kripke attacks the description theory of meaning for proper names, both as a semantic and a meta-semantic theory (though he doesn't use these terms). With regard to the latter sort of theory, he argues that people aren't in possession of any reference-fixing description for almost any proper name they use. In particular, using a description like 'whoever the person I got the name from is referring to' wouldn't do the job since that doesn't determine a reference on its own, relying as it does explicitly on the reference of the name as used by the provider. However, as substantive as Kripke's dispute with description theorists is, it's still the case that, with regard to my distinction between IMMs and DMMs, his account of proper names falls on the IMM side along with the description theories.

It seems clear to me that somewhere within our representational system of thought there have to be links with the world that do not themselves employ the contents of other representations. That is, some representations have to have DMMs if thought is to make contact with the world at all. So what are examples of DMMs? One example is the nomic-covariation account of concepts mentioned above. On this view, unlike reference-fixing versions (under which I would classify 'two-dimensional' accounts),[5] the content of a concept is determined by a lawful relation between a representation and the property it's about. While it may be that holding certain beliefs, or occupying certain other intentional states, plays a nomologically necessary role in maintaining the relevant lawful relation between the representation and its content, there is no appeal to the contents of these other states in determining the content of the target representation.[6] For instance, suppose it were the case that I wouldn't token [horse] in the presence of horses unless I held the background belief that my perceptual systems were generally reliable.[7] Still, it would be odd to say that the content of this belief was playing a meta-semantic role in determining the content of [horse]. Contrast this with a theory that says the way [horse] acquires its content is by its association with a description of the superficial properties of horses, together with the fact that it is these creatures, the horses, that satisfy the description. I want to call the latter mechanism an IMM, but the former a DMM.

Though we have found one direct meta-semantic link to the world, for concepts of properties and kinds, in order to secure singular reference we need a DMM for some singular concepts, or representations as well. I claimed above that Kripke's story about proper names doesn't count. However, noting the role of expressions like 'this baby' in baptisms, it's plausible that the items that play the foundational role for linking thought to the world, singular representations with DMMs, are

---

[5] See Chalmers, 2002.

[6] See Fodor, 1990, where he insists on this distinction between ways in which other intentional states help determine the content of a particular intentional state. He argues that this distinction allows him to consider his theory atomistic even while acknowledging the role of other intentional states in sustaining the relevant nomic connection.

[7] I'm not saying this is the case, but one can imagine reasons for thinking it might be.

indexicals and demonstratives. But before we investigate that question, bringing us to one of the two main tasks of the paper, there is one more distinction to introduce.

2.5 Pure Indexicals versus Demonstratives

Kaplan, 1989a distinguishes between pure indexicals like 'I', 'here', and 'now', on the one hand, and demonstratives, like 'this', 'that', 'she', and 'you' on the other. Pure indexicals vary their reference as a function of context, as do demonstratives, but the former require nothing more than the context to determine their reference. If I say 'I', that's enough to determine who it refers to, as my uttering 'now' suffices for picking out the relevant moment.[8] However, demonstratives require 'demonstrations' to complete them. One must point, or do something similar, to determine which object is referred to by a demonstrative.

On Kaplan's view, demonstrations are not part of the content of a demonstrative expression, but do serve the meta-semantic function of fixing their reference. There is a problem here, though—one noticed by Kaplan (1989b)—and it is particularly evident if we're interested in demonstrative thoughts themselves and not their linguistic expression. Suppose while at a party I ask someone nearby me, 'do you know who that man is?' while pointing, but for some reason I point in the wrong direction. Who have I referred to with the expression 'that man'? While one might try to make the case that I've referred to the man I'm (mistakenly) pointing at, it's clear that this can't be the man I have in mind. After all, if it were, in what sense would I have been making a mistake in where I pointed? So, at least for the thought behind the utterance, it's clear that the demonstration isn't really what determines reference. Kaplan takes this to show that it doesn't determine reference for the utterance either. Instead, he argues that the determining factor is the 'directing intention'.

Now I assume that what Kaplan means by the 'directing intention' is roughly what I mean by the mental demonstrative. The utterance 'that fly' expresses the corresponding thought-constituent, and I can't think of a better candidate for the directing intention behind the utterance than the thought-constituent it's expressing. Kaplan himself doesn't then go on to ask how the object of the directing intention is determined, but that's the question I want to address in this paper.

Before we tackle that question, though, let's say a bit more about pure indexicals. What makes indexicals special? We've already mentioned the feature of changing reference as a function of context. Another feature, emphasized by Perry (1979) is their untranslatability into non-indexical expressions. No non-indexical means of designating the present moment, say, has the cognitive force of 'now'. But I am

---

[8] Actually there's a question of boundaries that's not settled by the mere utterance—for instance, by 'now' do I intend just the time span of the utterance, this hour, this decade? Still, the core from which the relevant span emanates is determined by the time of utterance, whereas the indeterminacy of 'that' on its own is not merely a matter of fixing boundaries around a determined core.

interested right now in a different feature, the one that earned them originally the name 'token reflexive' (as Kaplan reminds us they were called by Reichenbach), and is presumably what underlies the other two features. A token reflexive is a representational element which refers by virtue of the context of its own tokening. The thinker of a token of 'I' is its referent, the location of a token of 'here' is its referent, and the moment of a token of 'now' is its referent. If one wants, one could capture what all of these token reflexives have in common by appending 'this' to thinker, location, and moment.[9]

In the cases just cited, 'I', 'here', and 'now', context determines reference by utilizing a feature of the tokening itself, its subject and spatio-temporal location. So is indexical reference achieved via an IMM or a DMM? It seems clear that this is a DMM, because I don't see how one can fashion an IMM that wouldn't require an indexical expression in the specification of the conditions, thus making it viciously circular and incapable of securing the term's reference. This is the upshot of an indexical's untranslatability in non-indexical terms. True, pure indexicals have character, and this may seem like an IMM. However, the character just serves to pick out the relevant feature of the tokening of the representation—whether it be its subject, location, or time—but that it is *this* particular token representation whose subject, location, or time is in question, this can't be expressed any other way than as 'this one'. What attaches it to itself has to be a non-intentional mechanism.

Let's push this a bit further. What could make the reflexive element in indexicals, the 'this' in 'this subject/location/moment', self-referential? We've seen that it can't be a matter of intention or content, as that appeal is circular. It isn't a matter of what causes the tokening, or any nomic connection between tokenings of this expression and, well, and what? Itself? What sort of law is that? Anyway, it's clear that if there is reference by nomic connection, it applies to general terms and properties/kinds, not individuals.

It's got to be then a matter of functional role. The role these terms play in inference, deliberation, and action determines the reflexive interpretation for them. I think the best comparison here is to logical operators. A standard line concerning the interpretation of the logical operators is that their conceptual/inferential role determines their interpretation. The fact that one is disposed to infer 'P' from 'P&Q' is what determines, in part, that '&' means conjunction. Similarly, the fact that 'I'-thoughts, 'here'-thoughts, and 'now'-thoughts play unique roles concerning the acquisition of perceptual knowledge, deliberation, and the execution of intentions to act, is what determines that they mean the subject, location, and time of their own tokening.[10]

---

[9] But isn't 'this' a demonstrative, so aren't we collapsing the distinction between so-called pure indexicals and demonstratives? No, what this shows is that 'this' when used token-reflexively—as in 'this moment', or even 'this sentence is false'—functions quite differently from when it's used demonstratively, as in 'this book'.

[10] Notice here that I use the term 'determines' and not 'constitutes'. This is the difference between appealing to functional role in one's meta-semantics and appealing to it in one's semantics. I'm

We now have at least one DMM for singular thought. However, so far we only have a direct linkage to oneself and one's spatio-temporal context. We also think about things other than ourselves, things that move through space and time. It is here that the special role of demonstratives, especially perceptually-based demonstratives, seems to come into play. We already saw this with respect to Kripke's use of baptisms to ground the reference of proper names. Another example that helps to illustrate the fundamental role played by demonstratives is the famous man with the martini, Donnellan's (1966) case for there being a referential use of definite descriptions. In that example, a person asks a friend at a party, 'who is that man in the corner drinking a martini?', but it turns out it's water, not a martini, in the man's glass. Donnellan claims that even though the man doesn't satisfy the description, the speaker's utterance still refers to him. Why? Well, he argues, it's because it was that man that the speaker had in mind.

Let's leave aside for the moment whether Donnellan is right that the utterance refers to the man.[11] What does it mean for him to have that guy 'in mind'? In the case where the speaker sees the man in the corner and mistakenly takes his drink to be a martini, it seems clear that his having the guy in mind is a matter of his demonstratively referring to him. How else would he be able to think of him in that case? So it seems clear that what provides our most immediate, direct link to concrete objects is demonstrative representation. The question before us is this: is the meta-semantic mechanism by which demonstrative thoughts refer to their objects as direct as the Donnellan case makes it seem, that we somehow directly have the object in mind? Does this mean that mental demonstratives possess DMMs? This is the question to which we now turn.

3. The Meta–Semantics of Token- and Type Demonstratives[12]

3.1 Perception, Attention and Demonstration

Let's return to my spotting the fly on the wall and thinking [that fly is annoying]. How does my thought [that fly] make contact with the fly? Two elements stand out

---

not saying that functional role is constitutive of meaning, or semantic value. On the contrary, '&' means conjunction, a certain truth function, period. Similarly, 'I' means me when I use it, period. Functional role only comes in to supply the meta-semantic mechanism by which '&' and 'I' acquire their semantic values. By restricting her doctrine to meta-semantics, the functional role holist can avoid at least one powerful objection to her doctrine, namely that holism entails that no two individuals ever share beliefs. The idea is that two individuals can differ significantly in their beliefs and yet meta-semantically map those beliefs onto the same contents for many of them.

[11] Kripke (1977) persuasively argues that Donnellan is confusing 'semantic reference' with 'speaker reference', and that there is no good argument that the utterance refers to the man, even though, in some sense, the speaker does. But my concern is with speaker reference here anyway.

[12] A precursor of the view developed here appeared in Levine, 1988.

in this case: visual perception and attention. The fly I'm demonstrating in thought is the fly that I see. If my eyes were closed (supposing I couldn't hear it either), I wouldn't know it's there and wouldn't be able to demonstrate it. Obviously visual perception is what makes the fly available for me to think about. But while perceiving the fly, in this case, is necessary for demonstrating it in thought, it's clearly not sufficient. After all, at the same moment I perceive the fly I also perceive a host of other objects.[13] What's needed in addition to perceiving the fly is *attending* to it.

Attention, especially visual attention, has been intensively investigated over the years by cognitive psychologists.[14] While a lot is known about it, it's still very difficult to say exactly what kind of process it is. Often the spotlight metaphor is used, though just as often researchers object to it. What seems consistent with the available research is to think of attention as a mechanism of selection, a way of making the selected object available for central processing of some sort. The analogy I like is between attention and the 'select' function in Word; highlighting the relevant text is a way for the program to select it for further processing, like moving, deleting, etc. Let's think of visual attention as doing something similar for thought; highlighting (with a spotlight maybe?) some object in (or portion of) the visual field for further mental processing.

Do we have all the elements necessary for our model of mental token-demonstrative representation? Not quite. Remember how Kaplan distinguishes between pure indexicals, like 'I' and 'here', and demonstratives, like 'that', on the basis that context alone suffices to provide a reference for the former, whereas some act, a demonstration, is required as well for the latter. It's just my uttering 'I'm hungry' that is enough to determine whose hunger makes the utterance true. But uttering 'that' in a room with a fly isn't enough to determine the referent; I need to point, or in some other way act to pick it out. Similarly, in thought, thinking [I'm hungry] is sufficient to determine whose hunger is at issue, and sufficient as well, together with the requisite other mental states, to get me over to the refrigerator. But thinking [that fly is annoying] requires a further act on my part to provide a referent for the mental demonstrative, a mental demonstration, a mental pointing. The question is, can attention fill this role? Or is some extra act of mental demonstration necessary as well; and, if so, how does it relate to attention?

In fact, I do think a separate mental act of demonstration is necessary. The main reason is that we seem able to attend to more than one object at a time.[15] Selecting an item makes it available for processing of various sorts, but then one has to go

---

[13] While there is an industry in vision science showing how little we actually perceive at any one time, despite our sense of a very rich and full visual field, it is still the case that we perceive more than one object at a time. See Pylyshyn, 2001, whose view I will discuss below.

[14] Pashler, 1998 provides a fairly comprehensive review of the literature up to that point.

[15] Pylyshyn's, 2001 and 2007 MOT (multiple object tracking) experiments make this vividly clear, but there is also other research that shows that divided, even cross-modal simultaneous attention is possible too. See Pashler, 1998, chapter 3.

and process it. In terms of the Word analogy, highlighting the text is one thing, putting one's cursor on it and moving it is another. So I think it's necessary to posit a kind of mental pointing—mental demonstrating—that serves to connect the demonstrative constituent in a structured thought (like [that fly is annoying]) with some object of attention. Employing the computer analogy again, think of the pointers to data structures. The demonstrative constituents in thoughts can thus be thought of as simply pointers to previously selected objects.

Once demonstration, mental pointing, has been introduced as an operation distinct from attention, one might wonder why we need attention anymore as part of the account. That is, why not just posit a pointing relation between a mental demonstrative and its referent and leave attention out of it? In the end this is an empirical question, and I admit to engaging here in some speculative psychology. But it does seem to me that there is an important role for attention to play here.

One might be able to imagine a cognitive architecture in which all perceived objects are available for demonstration simultaneously, and one just mentally points at one, as opposed to another. There might be nothing like selective attention in this architecture. But both introspection and psychological research makes clear that selective attention is a major aspect of how information from perception reaches more central cognitive processing in us. While I find I can attend to two objects while explicitly thinking [that one] of one of them, I don't find I can think [that one] of anything I'm not attending to. It certainly seems to be the case that we need to select some portion of what's perceptually available in order to think about it. The thought is still distinct from the selection, but it's the selection that, again, makes the object available to be thought about.

So far, then, the theory of mental demonstratives has the following form. A demonstrative thought involves a (structured) mental representation, one of whose constituents is a mental demonstrative, a kind of mental pointer. Mental demonstratives are directly referential semantic mechanisms, which means that their semantic values are the objects in the world they refer to; [that fly] directly refers to the fly. Mental demonstratives also have a Kaplanian character similar to that of public language demonstratives. That character, not itself explicitly represented, but presumably, like the case of pure indexicals, determined by functional role, embodies the rule that the demonstrative takes as referent whatever is currently being mentally demonstrated, which means whichever object of attention is currently pointed to.

There is one important disanalogy between the cases of mental demonstratives and public language demonstratives. Remember Kaplan's puzzle about how to deal with demonstrations that went awry. For some reason I point at the wrong guy when asking 'who is that man?'. In this sort of case one might use the Kripke (1977) distinction between speaker and semantic reference and say that my utterance is about the guy who I actually pointed to, though I myself am referring to someone else. Or, following Kaplan, one might say that it's the directing intention behind the utterance that determines its reference. One can see a case for going either way, though I myself prefer Kaplan's. But no such issue can arise in thought. There's no

analogue, it seems to me, of pointing going awry, or of attending to the wrong object.[16] If I'm thinking [that] while attending to something, what would it mean for me to be meaning something else? I don't see how there could be a further directing intention for which this act of attention or pointing misfires. The point is, when my mind occupies the relevant functional states corresponding to attention and pointing, that constitutes my directing intention, what it is I have in mind. That's where the intentional buck stops.

3.2 IMM or DMM?

Having our basic architecture in place, it's time to address our main question: is the meta-semantic mechanism by which a mental demonstrative refers an IMM or a DMM? Given the general model of token-demonstratives sketched above, the question breaks down into two: what sort of mechanism is mental pointing itself, and what sort of mechanism is attention? Since it's attention that makes the object available to be demonstrated, if attention is an IMM, mental demonstration would be as well.

With respect to mental pointing itself, it seems pretty clear that no intentional element is introduced here. For one thing, intuitively one doesn't think something like [the fly I'm now attending to is annoying], but [*that* fly is annoying]. We certainly aren't aware of using any description to pick out the object of our thought. While this may not by itself be decisive—maybe unconsciously we are employing a reference-fixing description—it certainly provides a *prima facie* reason for thinking no such descriptive mechanism is at work. But more important, as with indexicals, it isn't clear one can find an appropriate description that doesn't employ a mental demonstrative itself, leading to vicious regress. After all, if I'm attending to more than one object, which one is it I'm intending to pick out? Well, *this* one. How else can I represent it to myself? So what does do the meta-semantic work? As with indexicals, I think we need to appeal to the functional role. The crucial point is that it's the mental demonstrative's *instantiating* the relevant functional role, not it's satisfying an explicitly entertained description of the role, by virtue of which it points to whatever it points to. As with indexicals, again, the elements of the role will involve connections between perception, inference, and action.

Let's turn now to selective attention. Does attention directly select distal objects, or does attention employ an intentionally mediated mechanism? If the former, then mental demonstration, by virtue of its reliance on attention, will have a DMM. However, if the latter, then, also by virtue of its reliance on attention, mental demonstration will have an IMM. I now want to argue that attention possesses an IMM, not a DMM, and therefore mental demonstratives themselves possess IMMs. They are not the direct interface with the world they might seem to be.

---

[16] To use Shoemaker's (1968) phrase, the thought is 'immune to error through misidentification'.

At first blush it might seem odd to think of attention as intentionally mediated. After all, for the same sort of reason given above with respect to pointing, it doesn't seem as if one can analyze the act of attending as the entertaining of some description. I agree. I am not arguing that attending is somehow a kind of describing. Rather, attending, as I understand it, is a primitive mental operation that will receive a functional account, as do all primitive mental operations. But employing a reference-fixing description is not the only way for a representation to possess an IMM. Any meta-semantic mechanism employing an operation that takes as its immediate operand another mental representation also counts as intentionally mediated; or, to be more precise, it does so as long as the manner in which the operation performs its meta-semantic function crucially involves the content of the mental representation on which it operates. I claim that attention is just such an operation. In particular, attention acts immediately on perceptual representations of objects and/or spatial regions. The idea is that the visual system delivers a representation of the surfaces, spatial relations, and objects in the subject's visual field, and then attention selects some of these for further processing—including singling out in thought through a mental demonstrative. Therefore any representations—mental demonstratives in particular—that obtain their referent through attention count as having IMMs.

There are two kinds of consideration that support my claim that attention operates immediately upon mental representations, crucially utilizing their contents to perform its function. First, this follows from RTM. Second, extant empirical theories of attention seem to entail it. With respect to RTM, the basic idea is that mental processes and operations are defined over mental representations. Inference, for instance, is a causal process that takes mental representations (the premises) as input and yields a mental representation (the conclusion) as output. While the operations themselves are implemented in causal mechanisms that are sensitive to physically realized syntactic features, the intentional contents of the representations are what make the process an inference. Reasoning that is, semantically speaking, directly about the world, is accomplished, according to RTM, by internal operations on mental representations whose semantic interpretations involve distal objects and properties.

Now attention is of course a very different kind of mental operation than inference, but insofar as it's a mental operation, it comes under the general framework of RTM. Therefore, it acts upon extra-mental objects only through immediately operating on mental representations of those objects. Just as when I infer something about the world by deriving an appropriate mental representation whose content is about the world, so too I attend to an object by standing in the relevant functional relation to a mental representation of that object. Just as with inference, the semantic content of attention, as it were, is directly about the object. But the mechanism by which attention to the object is achieved involves a representation of it.

This last point is important to emphasize, so let me elaborate on it. It is easy to confuse my claim about attention with the claim that we can only attend to our own mental representations; that we are, as the old worry goes, stuck behind a 'veil

of ideas'. But just as a representational account of perception need not entail that we perceive our own ideas, my model of attention has no such consequence either. The answer in the perception case is that *what* we perceive are extra-mental, distal objects; but the *way* we perceive them is by forming percepts of them. So too, on my view, *what* we attend to are distal objects—like that fly on the wall; but the *way* we do it is by standing in the relevant functional relation to a percept of the fly.

If attention, and therefore mental demonstration, involves an IMM, and if, as I argued above, singular reference must be ultimately grounded in a DMM, where do we find that ground-level link between object and mind? Clearly, if what I've said so far is right, it has to be found in the percept itself. I will have more to say about this presently, but for now let's just say that some version of a causal theory of perceptual representation must do the job here. So my thought [that fly] gets to be about the fly because it points to a percept selected by attention, and that percept is about the fly. We can think of the semantics of mental demonstratives in Kaplanian terms. The character is a rule (not represented, but embodied functionally) to the effect that the mental demonstrative takes as its content whatever is the content of the percept to which it's pointing; the content, then, is whatever object that percept picks out.

Before turning to the empirical considerations, I want to clarify my argument from the nature of RTM by noting an ambiguity in the notion of representation itself. Understanding this ambiguity will help illustrate the sense in which mental operations require mental representations for their operands. When we say 'X represents Y', we can mean one of two things, depending on whether X is a cognitive subject or a mental representation. If the former, we are speaking of a mental act, and prior representation is required. But if the latter, no explicit mental operation, or act, is involved, and therefore the relation can hold directly between the representation and what's represented. Neglect of this ambiguity led Ryleans to object to mental representation theories of behavior on the grounds that representing such-and-such is itself a mental activity, and therefore requires, on that theory, yet further representations, leading to infinite regress.[17]

So, for instance, on the Jewish holiday Passover it is traditional to have a ritual meal, a 'seder', in which lots of symbolism takes place. In particular, bitter herbs are set out on the table to represent the suffering of the ancient Israelite slaves in Egypt. Now, when I engage in this practice, I am intentionally representing the suffering with the bitter herbs. This is a mental act on my part, and it clearly relies upon prior mental representation to carry it off. That is, I have to already have available to me a representation of both the bitter herbs and the suffering, in order to actively represent the latter by the former.

But when it comes to the primitive terms in my system of mental representation, no such operation, or act of representation is required. This is why primitive

---

[17] Ryle's (1949) own attack on the 'intellectualist legend', though not specifically framed in terms of RTM, had a similar form, and embodied a similar confusion.

predicates (say [water], as on Fodor's theory) can have DMMs. The mind doesn't represent water the way I use bitter herbs to represent suffering during the Passover seder; rather, by virtue of its symbol [water] standing in the appropriate nomic relation to water (whatever that is, I don't pretend to know), the symbol represents water. In a sense, this primitive representation is something that happens to the mind, not something it does. Because it's not something it does, it doesn't need further representations to carry it out, and so no regress is threatened.

3.3 Empirical Considerations—Pylyshyn's theory

Visual attention raises a host of issues that are studied by vision scientists.[18] What concerns me here, though, is the following question: what are the objects of attention? Opinion divides on this issue between those who take spatial regions to be the primary objects of attention and those who take it to be concrete objects (though some allow both). On Triesman's (Triesman and Gelade, 1980) 'feature integration theory', the visual system contains many 'feature maps', representations of the three-dimensional space around the subject through which features like colors and shapes are assigned locations. Attention to a location is what pulls features previously assigned to that location together into a single representation. For instance, having assigned green to a location on the color map, and square to a corresponding region on the shape map, attention allows me to see a green square in the relevant region.

How, on this theory, do I attend to a location? Well, the idea is that there is a master map that is connected to all the feature maps. When the 'spotlight' of attention is shined on a certain location on the master map, it pulls all the featural information for that location from the disparate feature maps. I can then see a unified object, a green square, say. Clearly, on this view, attention is directed on a region of space by way of an operation on a mental representation of that space, a perceptual–cognitive map. Of course, when you think about it, how else could it work? That is, assuming we do attend to locations/regions (an assumption Pylyshyn disputes), how else but with a mental map could we do it?

There is a lot of experimental evidence that seems to support the claim that we not only attend to locations, but also to objects.[19] What's more, according to Pylyshyn (2007) we only attend to objects. I want to look at his theory in a bit of detail because on the surface it appears to contradict my claim that attention acts immediately on representations. I will argue that if we look deeper, ignoring what I think is some misleading rhetoric about the theory, it actually supports my position.

---

[18] Actually some of the pioneering work on attention focused on hearing, and, as mentioned earlier, there is interesting work on cross–modal attention. However, I'm going to restrict my discussion to vision for ease of exposition, especially since the vast majority of studies deal with visual attention.

[19] See Pashler, 1998, chapter 5, and Pylyshyn, as cited in text.

Pylyshyn (2001 and 2007) argues that already in early vision distal objects are represented, and that these objects become the targets of visual attention. For our purposes, there are two crucial aspects to his theory: first, that it is physical objects, not spatio-temporal locations, that are the objects of attention; and two, that these distal objects are visually represented by mere tags, with no defining qualitative or locative features. He calls these objects of visual attention 'visual objects', and the mental items representing them 'FINSTs', for 'fingers of instantiation'[20].

To give the flavor of what he has in mind, consider the visual tracking experiments that provide a good deal of the empirical support for the theory. In these experiments, the subject is shown a computer screen with, say, 10 or 12 black squares on it. Shortly after the screen appears, the subject is presented with a cue, in the form of 3 or 4 of the squares lighting up, and then, after the highlighting ceases, all the squares start moving randomly around the screen for a short period. The subject's task is to attend to the squares that lit up and track their movements, being prepared to identify them after the motion ceases. Subjects are extremely good at keeping track of up to 4 or 5 squares at a time. Pylyshyn's hypothesis is that the visual system has a small set (4 or 5) of these visual object tags, or FINSTs, and it is through this form of representation that the subject keeps track of the objects on the screen.

Pylyshyn employs several different metaphors to capture the relation between the visual tags, or FINSTs, and external physical objects. Interestingly, these metaphors pull in opposite directions. On the one hand, the term 'FINST' itself connotes the mind reaching out and 'touching' the objects it's keeping track of. He also compares visual tags, or FINSTS, to demonstratives, claiming that they directly 'select' objects. This way of looking at the operation of visual tracking certainly seems to go against the model I'm advocating. On the other hand, he also speaks of objects 'grabbing' FINSTS, which connotes the more passive notion of representation I discussed above. Furthermore, he seems to advocate a causal theory of representation for FINSTS, which again fits well with my position. So what's the right way to view his theory?

I'm inclined to take the 'grabbing' metaphor, the idea that objects grab FINSTs, more seriously than the 'touching' metaphor, the idea that FINSTs 'reach out' to the world and directly select objects. Here's how I understand the way grabbing works on Pylyshyn's theory. An object presents a certain collection of spatio-temporally distributed features, such as shape, color, and motion information, which are detected by the relevant mechanisms in the very early visual system. When this happens, when the right sort of spatio-temporal feature distribution is detected, a FINST is automatically deployed. The object instantiating the relevant features is then the referent of that FINST. What's crucial is that none of this featural information is represented within the FINST, though it plays a causal

---

[20] The idea behind the term is that the tags are like extended fingers that touch objects and maintain contact with them as they move through space.

role in bringing about the FINST's tokening. On this picture, the causal path that begins with the object and terminates with the introduction of the FINST constitutes a straightforward DMM. Up to this point, the deployment of a FINST, all the processes count as pre-attentive. Attention, then, operates on the FINSTs. This makes sense of the FINST account of what's going on in the multiple object tracking experiments, and fits quite well with my model of demonstrative thought.

On the other hand, as I said, Pylyshyn also compares FINSTs to demonstratives, and talks of FINSTs as mechanisms for selecting objects. On my view, FINSTs, having DMM's, cannot themselves be demonstratives, and cannot select objects. However, I think these remarks of his can be explained in a way that is not really inconsistent with my position. To begin with, when it comes to the comparison to demonstratives, I think what's essential to the comparison for Pylyshyn is that FINSTs, like demonstratives, do not contain either descriptive semantic content or descriptive meta-semantic content—that is, they are semantically directly referential and they do not employ reference-fixing descriptions as meta-semantic mechanisms. In this respect, I too agree that FINSTs resemble demonstratives. It's just that I want to insist that the meta-semantic mechanisms by which mental demonstratives attach to their referents are mediated through visual percepts—indeed, if Pylyshyn's theory is right, through FINSTs—while obviously the same cannot be said of FINSTs themselves. I see nothing in Pylyshyn's comparison between FINSTs and demonstratives that need conflict with this point of contrast between them.

One might think that Pylyshyn's talk of direct selection of objects, however, is more difficult to square with my position. But here too I think appearances are deceiving. The contrast that concerns Pylyshyn is between an account that takes distal objects to be what visual attention (selection) is directed upon and an account that takes spatial regions to be what attention is directed upon, as in Treisman's 'feature-integration' theory. On the latter theory, as we saw above, visual attention is directed upon certain spatial regions, and then by virtue of detecting the right sort of coinstantiation of features, the visual system infers the existence of an object. For Pylyshyn, though detection of certain features coinstantiated in a spatial region by preattentive mechanisms may be causally necessary for the representation of an object in that region, attention itself is directed upon the object, not the region in which it is located.

While this contrast may sound like it violates my principle that attention, as a mental act or process, must operate upon a mental representation, in fact it doesn't. The point is that the contrast between Pylyshyn's view and the views he opposes can be easily reconstructed in my terms, and in a way that I believe preserves the essence of Pylyshyn's view. As I interpret the dispute, what both sides can admit is that attention immediately acts upon a mental representation. What divides them is what that mental representation represents: is it a spatial region or an object? FINSTs, after all, are the result of preattentive processes. Once formed, they can then serve as the domain for attention to act upon, and as attention is an intentional

mechanism, its semantic object — what it is that is attended to — is whatever it is that constitutes the semantic content of the objects in its domain of operation. So, while his opponents think that representations of location, constituents of an internal map of some kind, are the objects within the domain of attention's operation, Pylyshyn argues that it is FINSTs that serve that function. While on this view it is certainly correct to say that one is visually attending to, selecting, an object, one does this by standing in the relevant computational relation to a FINST. Thus the idea that mental processes always operate on mental representations is not in conflict with the idea that visual attention applies to objects.

3.4 Type–Demonstratives

Let's turn now to mental type-demonstratives. Suppose I'm in the midst of deciding what color to paint my house. I'm driving along and idly looking at houses to get ideas and pass one that strikes me as just what I want. I think to myself, looking at the house, [that's the color I want for my house]. Obviously I'm not interested in literally taking the paint off that house and putting it on mine. Rather, I want to put a token of that very same type on my house. [That color] in that thought, unlike the occurrence of [that fly] in [that fly is annoying], serves to demonstrate a type of color, not any token object.

Earlier I distinguished true type-demonstratives from pseudo ones. Pseudo type-demonstratives, remember, are really descriptions that contain token-demonstratives as constituents, as in [the color of that house]. So pseudo type-demonstratives bring up no new issues. The question, then, is whether anything like the model just developed for token-demonstratives can work for type-demonstratives as well. In fact, I think it can.

If mental demonstratives were pointers directly out into the world, then it would indeed be difficult to see how anything but individual objects could be demonstrated. Perhaps one could invoke a kind of Russellian acquaintance with universals, but it certainly seems harder to swallow than direct selection of concrete objects. Be that as it may, on the analysis of mental demonstratives developed above, they are pointers to mental representations. So the percept of the fly is incorporated into my thought about the fly by virtue of there being a constituent of the thought that points to the percept, the semantic content of which is the object of that percept. It seems plausible, then, to extend this analysis to type-demonstratives. That is, a mental type-demonstrative is one that points to a perceptual representation of a feature of an object, as opposed to the object itself, and takes as its semantic content the property that representation stands for.

The idea is that a percept represents an object as instantiating many visible properties. The percept might have the form of a FINST to which visual feature-predicates are attached, such as color, shape, location, etc. To demonstrate a feature type, one's mental demonstrative points not to the FINST, but to the relevant feature-predicate, the embodiment of one's perceptual experience of the feature in question. The demonstrative itself, though pointing at a predicate-like

representation, still plays the role of a singular term in thought. So the result is a thought about a demonstrated property.[21]

3.5 Comparison with Campbell's View

Campbell (2002) develops a theory of demonstrative reference in some ways quite similar to the one developed here, but also different in significant ways. I can't begin to do his account justice in all details, so let me focus on a couple of points of contact and contrast. To begin with, where Campbell and I seem to agree is in the central role played by attention. He argues that it is (conscious) attention that enables one to know what the referent of a demonstrative is. Clearly we are on the same page here, as on my view attention is what makes the object available for demonstration—and thus accounts for how we know what it is we're demonstrating.

There are two points of contrast, however: one not clearly a disagreement, but more a matter of emphasis, and the other a straightforward disagreement. The first point of contrast has to do with the role of consciousness. Campbell is insistent that what matters for knowledge of demonstrative reference is *conscious* attention. He contrasts how I know which object is referred to when someone utters a demonstrative expression and I can see where she is pointing with how I don't know which object is referred to if I only visually detect it through blindsight. I agree that there is an intuitive difference here, and certainly all of my examples involved conscious thoughts and conscious attention.

But I don't see any reason to restrict my account to conscious thoughts and attention. True, as Campbell's contrast between normal vision and blindsight makes clear, I can't *consciously* think of that fly without *consciously* attending to it. But that's because I've restricted myself from the start to conscious thought. If one is concerned primarily with the personal level—what the person is thinking and attending to—as Campbell certainly is, then, on the reasonable assumption that what's conscious and what's at the personal level are pretty much the same thing, it just falls out that it is only conscious attention that can do the job. I want to leave it open that demonstrative reference can take place at sub-personal levels, and therefore allow for the possibility of unconscious attention and demonstration. I don't know if this is possible, but I presume that it is an empirical question.

The other point of difference is more significant. Campbell adopts what he calls a 'relational view' of conscious attention, which he explicitly contrasts with the representational view. So on his view conscious attention is a direct relation between

---

[21] Notice, by the way, that it's not necessary that the type–demonstrative involve any reference to the object instantiating the feature, as a direct reference to the tokening of the feature within the requisite spatio-temporal regions should suffice. After all, I can look at the house, think 'that color is the one I want', and not really notice that it happens to be on a house. I can bypass the token-demonstration of the object entirely.

subject and distal object. While Campbell is of course willing to acknowledge the role of computation at the sub-personal level, he rejects RTM as a framework for thought at the personal level. For him, conscious experience has extra-mental entities directly in its content. It's not just that conscious experience can represent distal objects; everyone agrees to that. But rather, the object is literally a part of what it is to be that experience.

It would take me too far afield to embark on a defense of RTM, and thus argue directly against Campbell's relational view. For now I just want to note the difference between us, and to note in particular how this difference reinforces my earlier claim that adoption of RTM and holding that selective attention involves an IMM go together. I do want to make one comment though about his argument. An idea that keeps popping up in his discussion is that we need an account of how we can think about a mind-independent world, or how we can acquire a conception of such a world. He appeals to conscious experience to answer that question, and claims that it is only on the relational conception that it can perform that function. That is, it is by virtue of the fact that experience embeds distal objects within itself that we are able to acquire a conception of a mind-independent world.[22]

However, on my understanding of RTM, this sort of question doesn't really arise. Our thoughts aren't grasped by us, and no special account of how we understand concepts like that of an objective, mind-independent world is needed. Rather, we just think with these concepts, and that means we token the relevant mental representations. This is just a reiteration of my point earlier concerning the basic, more passive notion of representation. There are of course two questions that do arise in this connection. First, are the representations in question acquired, and if so, how? Second, what makes them about what they're about? The latter question is the meta-semantic question with which I've been concerned in this paper. The former question I haven't addressed, but I'm inclined to the view that at least certain basic concepts—embodied in the relevant mental representations—are innate, and so aren't acquired. If our basic concept of a concrete, extra-mental object isn't acquired, then the question of how we come by such a conception doesn't arise.

4. The Misuses of Demonstrative Concepts

In this section I want to consider, in light of the account of type-demonstratives just presented, certain ways that philosophers have recently appealed to them to solve certain problems. I have three particular cases in mind, and all of them share a common core feature: type-demonstratives are exploited as mechanisms of concept acquisition (or formation—the difference, if there is one, won't matter for what follows). I will argue that all three attempts to solve the problems they address encounter difficulties (to different degrees) that stem from the fact that

---

[22] See especially the discussion in chapter 6, section 3.

type-demonstratives cannot serve the function of concept acquisition/formation (or at least not in the sense required). I'll begin by briefly outlining the three cases.

4.1 The Three Cases

First, consider how McDowell (1996) employs type-demonstratives. McDowell wants to argue that in order for perceptual experience to play the justificatory role it clearly plays in fixing empirical beliefs, it's necessary that its content be fully conceptual. Only states with conceptual content can stand in a confirmatory relation to beliefs. However, there are arguments, going back to Evans, 1982, that perceptual experience cannot be attributed a fully conceptual content. In particular, there's the problem that perceptual experience seems to be richer, more fine-grained than conceptual contents are. For instance, we can make many finer color discriminations perceptually than we can form concepts of. Thus, when seeing that particular house color, in all its particularity, I'm not capable of forming a concept of that precise color, though of course I can form more generic concepts like red, or even scarlet. However precise my vocabulary is, though, it is never sufficient to capture the full range of possible discriminable color experiences.

McDowell's reply is to say that we can indeed form a concept of that particular shade of red. When looking at it I can think, 'that shade of red', and using this type-demonstrative, form a demonstrative concept of the relevant shade. The idea is that the content of my visual experience, that it attributes that shade of red to the surface of the house, is itself thereby conceptualized. Prior to the demonstration there is no representational content to the experience, which is why it can't play the confirmatory role regarding my belief that the house is that shade of red. However, through the formation of the demonstrative concept 'that shade of red', the experience enters the realm of the conceptual and representational.

Second, type-demonstratives are invoked by some, particularly Loar (1997) who appeal to the special nature of phenomenal concepts as a way of addressing the problem of the explanatory gap.[23] If indeed phenomenal properties are identical to physical or functional properties, why does it seem intelligible to ask why instantiating the relevant physical property should feel the way it does, and why do we feel that nothing about the physical characterization of the property helps to answer the question?

The response is to note that in situations where we feel gripped by this question, we are conceiving of the property (or state) in question under two very different concepts: from the first-person perspective, under what Loar calls a 'phenomenal concept' (which means a particular type of concept of a phenomenal state, not a concept that is itself phenomenal), and from the third-person perspective, under

---

[23] Perry (2001) also makes a similar move, though he doesn't explicitly address the explanatory gap, but rather the conceivability argument. However, his response to that argument could be used to address the gap as well.

the relevant theoretical concept. To explain how the two concepts indeed pick out the very same property in a satisfying way would require an integration of the two types of concepts that is impossible. We just aren't capable of seeing why a state we think of in this first-person way, using a phenomenal concept, should be a state we occupy when we occupy a state that is characterized in the relevant theoretical manner.

Crucial to the account of phenomenal concepts that underlies this claim of cognitive incommensurability is the idea that phenomenal concepts are 'recognitional' concepts that are expressed through type-demonstratives. We just recognize qualitative states by their feel, nothing more, and then reidentify them with phrases like, 'that one', and 'that again'. Sometimes Loar speaks of phenomenal concepts possessing modes of presentation that are constituted by the relevant phenomenal properties themselves, and the concept is then formed through a type-demonstrative. In this way, just as I might form a concept of that particular shade when recognizing it on the house, I do something similar when mentally demonstrating my own qualitative state.

Third, type-demonstratives have been employed in some theories of concept acquisition. For instance, in the standard Kripke-Putnam story about natural kind terms, we imagine the first use of a term like 'water' to have occurred at an initial dubbing in the presence of a sample of water. The speaker introducing the term says something like (or thinks, it doesn't matter), 'by "water" I mean this kind of stuff', pointing at the water sample as this is uttered. 'Water' is now about that stuff, and thus a new term is introduced into the vocabulary. A similar story could be told, presumably, about one's concept of water, the mental representation [water]. That is, the subject thinks something like [these samples are samples of this kind of thing], and now when one employs this concept it represents whatever kind, or type was demonstrated on that occasion.

4.2 The Problem

With respect to the first two philosophical uses of type-demonstratives just described, I think it's clear that they must involve true type-demonstratives, not the pseudo-type-demonstratives that are really definite descriptions with embedded token-demonstratives. McDowell himself distinguishes cases where you actually see the color and think 'that color' from cases where you specify the color indirectly, as in 'the color of the house we lived in ten years ago', when you can't actually see it or even bring an image to mind. Since he's after a way to conceptualize the content of current visual experience, of course only the first kind of case would concern him. But since the pseudo-demonstratives don't exploit one's current perceptual contact with the color, they clearly can't do the work he wants done.

But can even true type-demonstratives do this job? I don't see how. If we adopt the account I presented in section 3.4, a type-demonstrative is a mere pointer to a perceptual feature representation that serves to incorporate the content of the perceptual representation into the thought of which it is a constituent. Thus all

the representational work is done by the perceptual representation itself, with the type-demonstrative merely selecting it for the thought in question. In no way, on this account, can a type-demonstrative itself be seen as a concept of anything, or a way of conceptualizing anything. That work has to be done already before any significance, or content, can be attached to the demonstrative element. In particular, already in perception the feature, the shade of color, must have been represented.

While my objection to McDowell's proposal is easy to see in the context of my treatment of demonstratives, I don't think it's necessary to accept the account in all its details to find the main idea behind the objection cogent. Demonstrating is something we do, something the mind does, and in doing it we know what it is we're doing; in this case, what it is we're demonstrating. When you see a color and think 'that color', the seeing is prior to the demonstrating, or else you really don't know what you're demonstrating (as in the case of pseudo-type-demonstration, where you really don't know which color you've picked out). But if the seeing, the perceptual experience, is prior to the demonstrating, then the demonstrating can't be what captures, or brings into existence, the content of that experience.[24]

Of course one might say this in response. True, the percept already represents the color and the demonstrative picks out the color by virtue of its relation to the color representation within the percept. But within the percept the representation is non-conceptual, and what turns our representation of the color into a conceptual one is the demonstrative link that is embedded in a conceptualized thought. Perhaps that's right. I personally find the conceptual/non-conceptual distinction extremely obscure, meaning different things in different discussions, so I take no position here on the controversy that McDowell is concerned with. Still, it's clear that McDowell, who wants to deny any representational content to perceptual experience unless it's conceptualized, cannot make use of this response, since it explicitly grants the percept a non-conceptual content. The point is, no new content is generated by the type-demonstration itself, so McDowell's ploy to get around the Evans argument won't work.

I think it's also clear that what Loar and company have in mind—or should have in mind—when they characterize phenomenal concepts as involving type-demonstratives are true type-demonstratives, not the pseudo ones. The idea is supposed to be that one picks out the phenomenal type directly by virtue of experiencing it, not indirectly, descriptively, as whatever phenomenal type it was that that experience tokened. In fact, it seems that the proposal for grounding a phenomenal concept in a type-demonstrative is almost exactly like McDowell's proposal for grounding a color concept. Therefore, it shouldn't come as a surprise that it stumbles on the same obstacle. One must already be aware of, and therefore

---

[24] Note that it doesn't help to adopt Campbell's relational view of experience, a view McDowell certainly shares. While it's true that on that view the content of experience somehow directly contains external objects, it doesn't help McDowell since experience already is supposedly conceptualized.

have represented, the phenomenal type before one can demonstrate it. The type-demonstrative itself can't be the primary means of representing the phenomenal type.

One might however try to distinguish the phenomenal concept move from the McDowell move as follows. When we are talking about objects or features out in the world, then of course one can't demonstrate them directly without first representing them. Mental demonstratives, as pointers, can only immediately 'touch' things inside the mind, and only through them refer to what's outside the mind.[25] Given that McDowell is talking about forming concepts of the features of external objects, his appeal to type-demonstratives can't work. But the phenomenal concept strategy is about forming concepts of mental states themselves, which, by virtue of already being inside the mind, can be pointed to, demonstrated, directly.

There are two reasons this response won't work. First, what's inside the mind is the token phenomenal experience, not the type. If we were dealing with a pseudo-type-demonstrative, perhaps then we could characterize the relation between the token-demonstrative constituent and the individual experience as unmediated by any representation of the experience, and then employ the descriptive apparatus, 'the phenomenal type of which this ('this' being the embedded token-demonstrative) is a token' to represent the type.

But, as mentioned above, a pseudo-type-demonstrative is not the appropriate mechanism for what Loar and company have in mind. If we're dealing with a true type-demonstrative, then how is the correct type to be identified—pointed to—unless it's already represented? After all, every token state is a token of an indefinite number of types. Which of these types is the type-demonstrative picking out? If its reference is determined along the lines I outlined in section 3.4, that's not a problem, since by virtue of its functional role it is interpreted to be referring to what is represented by the symbol to which it points. But if it's not pointing at a representation and borrowing its content, as it were, then the problem of serious ambiguity becomes insurmountable. The functional role of pointing alone can't determine which of the many properties the token pointed to instantiates is the property being picked out.

While the objection just presented suffices, to my mind, to render the response untenable, it seems to me a more basic point is at issue here. The principal theme of my entire discussion of demonstrative thought is that demonstration—and along with it, selective attention—qua mental process, acts upon mental representations. It's the content of what's immediately connected to the pointer that matters, not its non-intentional properties. So it's not to the point to say that since phenomenal states are already in the head they can be the direct objects of demonstration, of mental pointing, as if the problem with direct pointing at external objects was just their distance. The barrier to be overcome isn't merely physical, but of another sort.

---

[25] Louise Antony recently characterized this as the 'proxy argument', the idea that the mind requires proxies for external objects to act upon.

Pointing, demonstrating, as I've characterized it, is a mental operation, and by virtue of that fact it takes objects with contents as its operand. Therefore demonstrating must always be mediated by prior representation, even if what's being represented happens already to be inside the head.

Before leaving this discussion of the phenomenal concept strategy, I want to enter an important qualification of my critique. I do not take myself here to have presented an objection to the entire phenomenal concept strategy for dealing with the explanatory gap and the conceivability argument.[26] I think there are other ways of trying to characterize what is special about phenomenal concepts that don't rely upon the notion of a type-demonstrative, some of which may involve Loar's idea of a recognitional concept—a notion that I think is separable from the notion of a type-demonstrative. My only point is that insofar as one employs the notion of a type-demonstrative to characterize what is special about phenomenal concepts, it can't do the work it's being asked to do. The crucial representational function must already be performed prior to the employment of any demonstration.

Finally, let's briefly consider the idea that one can acquire concepts of kinds, like water, by dubbing ceremonies similar to the sort Kripke envisioned for names of individuals. Unlike the McDowell and phenomenal concept cases, it's not obvious that what people have in mind here involves true type-demonstratives. Again, as mentioned above, the story goes like this. A person points at a pool of water and says something like 'By "water" I mean this kind of thing'. I think it's possible to see this either as a true type-demonstrative or as a pseudo-type-demonstrative. Either way, though, the account suffers from a serious problem.

Let's assume that what's supposed to be involved here is a true type-demonstrative. It should be obvious by now what the problem is. True type-demonstratives only refer to whatever is already represented. Therefore, to directly demonstrate the type, or kind, water, by a demonstrative, one must already possess the means for representing that kind. But then the demonstrative act can't possibly be the means of acquisition of the concept.[27]

It's a little more complicated if what's intended by this account of concept acquisition is a pseudo type-demonstrative. In this case, instead of literally '*this* kind of thing' what one means is 'the kind *of this* thing', where the demonstrative is of

---

[26] See Levine, 2007 for more general discussion of this strategy.

[27] Unless, of course, one makes use of a possible reply I mentioned briefly above in my discussion of McDowell. If one takes the conceptual/non-conceptual distinction seriously, it is consistent to claim that one acquires a conceptual representation of a kind by employing a type-demonstrative whose reference to the kind is secured through a non-conceptual representation of that kind. This couldn't work for McDowell since he doesn't allow non-conceptual representation, but if one does there is no argument from my account of demonstratives to undermine this move. However, it is still the case that for this strategy to work, the kind itself, and not merely some corresponding superficial sensible property, must be the content of the non-conceptual perceptual representation. I imagine many would find this claim to be implausible; and even if not, it's still the case that the original representation of the kind is not brought about by way of the demonstration.

the token variety picking out the pool of water one is looking at. Since one could have a percept of an object without determining its kind, the necessary conditions for the token-demonstrative to refer are met. From there the kind can be picked out by the definite description that contains the token-demonstrative.

The problem for this way of acquiring a concept is one we've seen already in our discussion of using pseudo type-demonstratives to form phenomenal concepts.[28] There is just too much ambiguity associated with the phrase 'the kind of this thing' or 'the type of which this thing is a token'. Individuals are members of indefinite numbers of kinds, tokens of indefinite numbers of types. If we don't already know which kind it is we're picking out, how does the description help to determine it? And if we do already know which kind we're picking out, then we already have the relevant concept, and what do we need the demonstrative for? Perhaps there is a way of overcoming the ambiguity problem for natural kinds, but I wouldn't bet on it. At any rate, it's clear that the notion of a true type-demonstrative can't help with this problem.

5. Conclusion

In this paper I have proposed a model of demonstrative thought that treats both token-demonstratives and type-demonstratives in a similar fashion. In both cases, the demonstrative element in the thought involves a pointer to the relevant perceptual representation, one made available by attention. I have argued that it follows from this model, as well as from independent considerations, that demonstration, as a mental act, operates directly on mental representations, not external objects. That is, though the relation between a demonstrative and the object or property demonstrated is semantically direct, it is meta-semantically mediated by a perceptual representation. Finally, I argued that once we recognize the essential role played by perceptual representations in establishing the link between demonstratives and their objects, we see that the notion of a 'demonstrative concept' cannot perform the various philosophical functions that have been attributed to it.

Department of Philosophy
University of Massachusetts

References

Campbell, J. 2002: *Reference and Consciousness*. Oxford: Oxford University Press.
Chalmers, D. 2002: The Components of Content. In D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

---

[28] Jerry Fodor has made this criticism in conversation.

Donnellan, K. 1966: Reference and definite descriptions. *The Philosophical Review*, 85, 281−304.

Dretske, F. 1981: *Knowledge and the Flow of Information*. Cambridge, MA: Bradford Books/MIT Press.

Evans, G. 1979: Reference and contingency. *The Monist*, 62, 161−89.

Evans, G. 1982: *The Varieties of Reference*. Oxford: Oxford University Press.

Fodor, J.A. 1975: *The Language of Thought*. New York: Thomas Crowell, and Cambridge, MA: Harvard University Press.

Fodor, J.A. 1990: *A Theory of Content and Other Essays*. Cambridge, MA: Bradford Books/MIT Press.

Kaplan, D. 1989a: Demonstratives. In J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*. New York: Oxford University Press.

Kaplan, D. 1989b: Afterthoughts. In J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*. New York: Oxford University Press.

Kripke, S. 1977: Speaker's reference and semantic reference. In P. French, T. Uehling Jr. and H. Wettstein (eds), *Studies in the Philosophy of Language*, Midwest Studies in Philosophy, no. 2. Minneapolis, MN: University of Minnesota Press.

Kripke, S. 1980: *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Levine, J. 1988: Demonstrating in mentalese. *Pacific Philosophical Quarterly*, 69, 222−40.

Levine, J. 2007: Phenomenal concepts and the materialist constraint. In T. Alter and S. Walter (eds), *Phenomenal Concepts and Phenomenal Knowledge*. New York: Oxford University Press.

Loar, B. 1997: Phenomenal states. In N. Block, O. Flanagan and G. Güzeldere (eds), *The Nature of Consciousness*. Cambridge, MA: MIT Press.

McDowell, J. 1996: *Mind and World*. Cambridge, MA: Harvard University Press.

Pashler, H. E. 1998: *The Psychology of Attention*. Cambridge, MA: Bradford Books/MIT Press.

Perry, J. 1979: The problem of the essential indexical. *Noûs*, 13, 3−21.

Perry, J. 2001: *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.

Pylyshyn, Z. 2001: *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: Bradford Books/MIT Press.

Pylyshyn, Z. 2007: *Things and Places: How the Mind Connects with the World*. Cambridge, MA: Bradford Books/MIT Press.

Ryle, G. 1949: *The Concept of Mind*. Chicago: The University of Chicago Press.

Shoemaker, S. 1968: Self-reference and self-awareness. *Journal of Philosophy*, 65, 555−67.

Treisman, A. and Gelade, G. 1980: A feature integration theory of attention. *Cognitive Psychology*, 12, 97−136.