DEMONSTRATIVE THOUGHT
Joseph Levine, University of Massachusetts, Amherst
January 20, 2007

1. Suppose a fly is buzzing around the room. I’ve noticed it and it’s bothering me. I locate the fly swatter and then visually search the room for the fly. I suddenly see a fly on the wall and I say to myself, “Aha, I’m gonna get that fly now!” I’ve just expressed a thought, and clearly 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 (md).

Demonstrative expressions are of two types: token and type. 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. “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)” - an md - let’s assume there is also a thought constituent - also an md - corresponding to the type-demonstrative “that color”. The significance of distinguishing these two types of md will emerge as we proceed.

I have two principal goals in this paper: First, I want to investigate how md’s function; in particular, I want to see how, and in what way, they afford us direct access to what is demonstrated. Second, I want to criticize the way that the notion of a “demonstrative concept” has been used to solve certain philosophical problems. While the discussion of the first question concerns both token- and type-demonstratives, the second topic principally concerns type-demonstratives. When philosophers speak of demonstrative concepts they usually have type-demonstratives in mind. However, since my critique of the uses to which this notion has been put is based on my answer to the first question, much of the discussion concerning token-demonstratives will have a bearing on this topic as well.

My plan for the paper is as follows. In section 2 I introduce some distinctions that will help structure the discussion to follow. In section 3 I address the first issue mentioned above concerning token-demonstratives, with special emphasis on the relation between mental demonstration and perception. In section 4 I apply the results of the discussion in the previous section to the case of type-demonstratives. Finally, in section 5 I look at various ways some philosophers have appealed to type-demonstratives to solve certain problems in recent philosophical literature. I argue that, in light of the account developed in section 4, the appeal to type-demonstratives can’t do the work it’s supposed to do.

2. 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 statement (or thought): what are its truth conditions, how does it behave in modal contexts, etc? The direct reference theory, considered as a semantic theory, tells us that for singular terms that refer directly, the referent itself is the content. A direct reference theory for general terms might say that the property or kind picked out by the predicate is its content. The idea behind direct reference as a semantic theory is that the mode of presentation, if there is one, is not itself part of the content, not part of what is literally asserted or thought.

The meta-semantic question concerns how the content of a term, whatever it turns out to be, got to be that term’s content. For instance, on Kripke’s (1980) causal theory of names, the semantic question is addressed by the claim that a name’s content is the object named, and the meta-semantic question is addressed by the idea of a historical chain of uses leading back to the relevant object. What it means to assert “Socrates is wise” is just to attribute wisdom to the guy, Socrates; nothing more to it. But there is a complicated process of acquiring and transmitting the name through history that is responsible for the fact that current uses of “Socrates” are about that particular historical figure.
In considering the metasemantics of singular terms, it’s helpful to employ, or rather, adapt, Donnellan’s (1966) famous distinction between attributive and referential uses of descriptions. I’m quite sympathetic to Kripke’s (1977) objection that when it comes to linguistic meaning - what he called “semantic reference”, as opposed to “speaker reference” - Donnellan is mistaken in thinking there is such a thing as a referential use of a definite description. Descriptions refer only by picking out whatever it is that satisfies the conditions specified in the description. But since thought is my topic here anyway, it’s speaker reference I really care about, and here it does seem as if there are these two basic mechanisms for establishing reference: refer to whatever satisfies certain descriptive conditions (the attributive mechanism) or refer directly (the referential mechanism). What a “referential mechanism” might be will be explored presently, but to start it seems intuitive that it has something to do with demonstrative ways of picking things out.

Donnellan of course didn’t mention demonstratives explicitly in presenting the referential use. But if we look at his cases - for instance, the infamous “man with the martini” - it seems clear that the reason the speaker is referring to that man even though he doesn’t meet the relevant descriptive condition (he’s drinking water, not a martini) is that he has available to him another means for picking him out, and it’s not another description. Rather, what makes the guy drinking water over there this speaker’s referent is that the speaker sees him and thereby demonstratively picks him out as “that guy”.

Suppose, then, we take Donnellan’s referential/attributive distinction to apply specifically to metasemantic mechanisms. An attributive metasemantic mechanism (AMM) is one that picks out an individual as the referent of a representation by virtue of its satisfying uniquely identifying conditions that are associated with the use of that representation - either because these conditions actually constitute the semantic content of the representation, or because they serve as the reference-fixing condition for it. So long as we have a metasemantic account for general terms and logical operators, then, using Russell’s theory of descriptions, we have a fairly complete story for how singular reference comes about via an AMM.

One can characterize a referential metasemantic mechanism (RMM) both negatively and positively. On the negative conception, an RMM is just any metasemantic mechanism that isn’t attributive. On the positive conception, however, the idea is to capture this sense that sometimes the mind has a direct hold on the object, or is in some way, to use Kaplan’s (1968) early formulation, “en rapport” with the object. It’s clear from the way Donnellan describes the referential/attributive distinction that he has something like this in mind. He speaks of thinking of a particular individual, as opposed to thinking of whoever it is that satisfies a particular condition, as if in some cases minds can just directly think of individuals without any conceptual or cognitive mediation. The idea of being en rapport is closely connected of course to Russell’s notion of being “acquainted” with an object.

Though I happen to think there is an important philosophical role for the notion of acquaintance to play, for the purposes of this paper I will stick to the negative conception of an RMM. Prima facie, to avoid circularity or infinite regress, all reference must be ultimately determined by an RMM. So even if one introduces a name by a purely descriptive reference-fixer, the metasemantic question arises for the terms, including the general terms, in the relevant description. Perhaps one can get away without specifying RMM’s for any singular terms, relying always on AMM’s, but one must still at least provide RMM’s for the general terms.

However, there are good reasons for insisting on there being RMM’s for at least some singular terms as well. For one thing, names, demonstratives, and indexicals just don’t seem to refer by way of reference-fixing descriptions. For another, there is reason to believe that without a directly referential mechanism for at least some singular (mental) terms, one couldn’t get a directly referential metasemantic mechanism for general terms either. That is, our semantic attachment to properties of objects plausibly relies on our initially picking out the objects instantiating those properties. Singular thought is plausibly the means by which we connect to the world in the first place.

1 Kaplan discusses this relation to demonstratives in his (1989b).

2 In Levine (1998 and 2001) I tried to capture what is basically this distinction in terms of “ascriptive” versus “non-ascriptive” modes of presentation.
Turning back to the example of Kripke’s theory of proper names, two points relevant to our discussion arise. First, the particular metasemantic mechanism Kripke describes, the historical chain of uses, doesn’t clearly fit into either the AMM or the RMM category. The problem is that though the historical chain is supposed to take the place of a descriptive reference-fixer, it isn’t as if no intentional element is involved. As Kripke himself admits, it’s crucial that in the transmission process there be some intention on the part of the recipient of the name to use it to refer to whoever is referred to by the donor (or the relevant community). Just hearing a sound and then using it - say as a name of one’s pet - doesn’t count as genuinely acquiring the name. Thus there is a descriptive component here, namely something like “whoever so-and-so means”.

On the other hand, unlike pure cases of AMM’s, this description can’t ground the reference of the name since it’s parasitic on whatever metasemantic mechanisms established the initial reference. What this shows is that there are mixed cases, terms in thought that get their reference through a combination of an RMM and an AMM. What matters most for our purpose here, though, is that the mechanism that ultimately grounds the term’s reference be either a pure AMM or a pure RMM.

The second point that emerges from Kripke’s account of proper names is that insofar as the relevant historical chain is grounded in an RMM - which, he argues, is the case for most proper names - the mechanism in question is a demonstrative thought. The way the story goes (and remember, for Kripke it’s just a “story”, not a full-fledged “theory”) there are two ways a name acquires its reference: by a “dubbing” ceremony or by a reference-fixing description. Dubbing is a matter of associating the name with an individual that is already picked out in thought in some other way. Well, which way? The only way I can imagine is demonstratively. I pick out some individual in thought as, say, “that baby”, and then associate the name with her. Presumably we’re talking about a perceptually-based demonstrative too. “That baby” is the one I’m currently perceptually attending to.

I doubt whether there are any candidates for singular thoughts with RMM’s that don’t involve demonstratives or indexicals. As we saw already, names bottom out in either demonstratives or pure AMM’s. But whether or not this is the case, what concerns me now is whether perceptually-based demonstratives, like my use of “that fly” in the original example, possess genuine RMM’s. That is, even in the case of demonstrative thought, does the mind directly refer to objects? What we need to do is dissect demonstrative thought into its components and trace the referential route back to the object. This is the task of section 3.

3. In “Demonstratives” (Kaplan 1989a), Kaplan distinguished between pure indexicals like “I”, “here”, and “now”, on the one hand, and demonstratives, “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. 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 metasemantic function of fixing their reference. There is a problem here, though, one noticed by Kaplan in his “Afterthoughts” (Kaplan 1989b), and it is particularly evident if we’re interested in demonstrative thoughts themselves and not their linguistic expression. Suppose when I’m asking who that guy drinking the martini is at the party I point instead and say,”do you know who that man is?”, 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

---

3 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 out 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.
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 (md). 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 now.

Before tackling demonstratives, however, let’s discuss the case of 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 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.4

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 RMM or an AMM? It seems clear that this is an RMM, because I don’t see how one can fashion an AMM 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 AMM. 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-attributive 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 - which presumably would count as an RMM - 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.5 In fact, it makes sense to see indexicals as really part

4 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”.

5 Notice here that I use the term “determines” and not “constitutes”. This is the difference between appealing to functional role in one’s metasemantics and appealing to it in one’s semantics. I’m 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 metasemantic mechanism by which “&” and “I” acquire their semantic values. By restricting her doctrine to metasemantics, 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
of the logical apparatus. Where “&” serves to mediate internal inferential relations among representations, the indexicals serve as hooks to attach the representation system to the world. The way these hooks work is to exploit their own spatio-temporal location (as well as the identity of the system of which they are parts).

So token-reflexivity is our first bona fide RMM. Are demonstratives token-reflexive in this way? That is, is there a feature of the tokening of an md that itself provides its referent? Certainly just looking at the immediate circumstances of the md itself - its spatio-temporal location, for instance - doesn’t seem to determine what it refers to. The fact that demonstratives are not token-reflexive in the way so-called “pure indexicals” are is of course why, as Kaplan put it, they require completion by a demonstration or directing intention. But since we are now considering how the referent of the directing intention is itself determined, we need to look elsewhere for the requisite completion.

Demonstration is always thought of as a kind of pointing, or even touching. The idea is supposed to be that through a demonstrative representation the subject gets “directly ahold” of the object, independently of any attributive conditions. So in the Donnellan case we characterize the speaker as “having that guy in mind,” and so it doesn’t matter that the descriptive conditions she expresses to communicate who she has in mind don’t actually fit the object of her thought. She has a more direct route to him. Again, it’s as if her mind extends a finger and touches him.6

Of course eventually the mind has to “touch” the world, because, as we mentioned above, AMM’s can’t be basic. There has to be some mechanism by which a purely, directly referential connection is established, in order for all other reference to get off the ground. Many discussions of this issue, as in Kripke’s causal “story” of names, take it for granted that demonstratives are the likely candidates for the mental “touchings”. But I think this can’t be actually. Once we think clearly about what it is to demonstrate, we’ll see that it already presumes referential contact with what’s demonstrated, and therefore can’t serve as that ground-level link.

Before making the argument, let’s step back and consider again our taxonomy of metasemantic mechanisms. The distinction drawn above between an AMM and an RMM has to do with whether reference is achieved through an object’s satisfying a description entertained by the referring subject or not. But there is a further distinction to be drawn, one that places AMM’s and some RMM’s on one side and other RMM’s on the other: metaphorically, the distinction is between a world-to-mind mechanism (WMM) and a mind-to-world mechanism (MWM).

Here’s what I have in mind. In the case of a pure AMM, it’s clear that the subject is referring to an object by laying down conditions it must satisfy, and whichever object satisfies it, that’s the object referred to. This mechanism counts as what I’m calling a MWM mechanism, since the mind imposes a condition on the object in the world to satisfy.

A good example of what I mean by a WMM mechanism is the information-based semantics for mental predicates espoused by philosophers like Dretske (1981) and Fodor (1990). A mental term H means “horse”, or “is a horse”, by virtue of the fact that H (or the property of tokening H) stands in the appropriate nomic relation to horses (or the property of being a horse). There is no intention to refer to horses, just a causal/nomic reflex in response to horsehood. After all, how could there be such an intention to refer to horses, without already possessing the means to refer to them? That is, an intention to refer to horses must itself contain a component that represents horses, so the original source, or ground of the referential relation cannot be the intention itself. In the realm of singular terms, indexicals, through being token-reflexives, would also count as WMM’s. The point is that the mere fact of where and when the token occurs is the ground of its reference, not any mental act.

Given the examples above, it’s tempting to think that the MWM/WMM distinction lines up with the AMM/RMM distinction. The mind reaches out and “grabs” an object in the world when it frames conditions that the object of its thought must satisfy, and the world “grabs” a mental constituent when an object occupies the appropriate link in the causal chain leading to that constituent’s tokening. However, it

---

6 Interestingly, the metaphor of touching is explicitly invoked by Pylyshyn (2001 and forthcoming) in his theory of visual object representation. I will discuss Pylyshyn’s theory in detail below.
is precisely with demonstratives that I think these two distinctions come apart. On the one hand, unlike straightforward cases of attributive reference-fixing, such as Evans’s (1979) “Julius” case, there do not seem to be any serious candidates for descriptive reference-fixers for demonstratives. On the other hand, unlike the other examples of RMM’s mentioned above, demonstrative reference does seem to involve an intentional act, demonstration, and in this sense it doesn’t seem as if an object in the world is grabbing a term, but rather as if the mind is reaching out and grabbing an object, thus fitting the MWM model better than the WMM model. Demonstratives, then, seem to involve a directly referential mechanism that works in the mind to world direction. How should we understand this?

Let’s look a little more closely at what mentally demonstrating an object comes to. To begin with, demonstrating seems to bear an intimate relation to perceptually attending; in fact, it wouldn’t seem implausible to me to just identify demonstrating with attending, though I don’t think we need insist on that. The point is, when I demonstrate that fly on the wall, visually attending to it seems to be an essential component of the process. Attention itself seems to be principally a matter of selection. This idea is captured by the psychologists’ “spotlight” metaphor for attention - the object of attention is that upon which the spotlight shines. Another analogy I like is the “select” function on a computer. To attend to an object is to highlight it, as we do with the text on a computer screen.

Seeing demonstration as a kind of mental selection, we get the following picture of a demonstrative thought. When I think [that fly is annoying me],9 the md constituent of the thought is a device for selecting the object to which the predicate applies. Semantically, as we’ve been assuming all along, the value delivered to the overall propositional content is just the object selected; in this case, the fly. Metasemantically, selection is not an AMM in that it doesn’t involve any sort of conceptualization, any specification of conditions to be satisfied in order to count as the selected object. So it appears to be a genuine RMM. But now we face the question, can one directly, in an unmediated way, select an external object? I think not, and therefore propose that though selection is not a conceptualized, attributive mechanism, it isn’t direct and unmediated either.

What sort of mediation might there be aside from an AMM? The mediation I have in mind is that provided by perception itself. The idea is that mental demonstration selects its object through a kind of deferred ostension.9 The md functions as a kind of pointer to the perceptual representation of the object, with the semantic value of whatever it is that the percept is a percept of. When I think [that fly], I have a current percept of the fly that is, functionally speaking, the immediate object of selection, and it is by virtue of my selecting that percept in thought that I, in this deferred manner, select the fly itself. The point is that my only access to the fly in thought is via my perception of it, and so I can’t really directly select it without going through my perceptual representation. However, it is not as if I entertain a description in thought of the form [whatever it is that that percept is a percept of]; rather, there is a deictic element in thought that points directly at the percept and is thereby assigned the semantic value of whatever object it is that this percept represents.

So why interpret the md mechanism in this way, as a mediated mechanism, rather than as a direct, unmediated relation to the object? Before giving my reasons for opting for the mediated mechanism, it’s worth elaborating on just what the import of this mediated/unmediated distinction is in this case. In fact, one might question whether there is any difference on the grounds that however one understands the metasemantic mechanism in this case, it’s clear that whatever causal route is going to connect the md to the object must travel through the perception of the object. After all, it is the fly that I’m looking at that I’m demonstrating.

Despite the fact that the causal route from the fly to the md about the fly must go through the visual percept of the fly on any story, there is still a principled distinction between treating the metasemantic mechanism of the md as a mediated one and treating it as unmediated. On the latter story,

---

7 From now on I will indicate mental representations with brackets, leaving quotation marks for natural language expressions.

8 If there is one - what to do about empty demonstration is an issue I can’t deal with here.

9 I proposed essentially this idea in Levine (1988) in the context of addressing a different issue.
what does the work of connecting the md [that fly] to the fly is the fly’s occupying the right link in the causal chain leading to the tokening of the md. That causal chain does go through the percept, of course, but the metasemantic mechanism does not exploit the semantic content of the percept to determine the md’s semantic content. The only role the percept plays is that of another link in the causal chain.

On the account I favor, however, the md’s semantic content is determined by the content of the relevant percept. The metasemantic mechanism involves a pointer to the percept whose content could be captured in the description “whatever it is that that percept represents”, though it gets this content by virtue of functional role, not an explicit description. So what we have is a mediated RMM; a purely causal relation to a percept, and, by virtue of that causal relation, inheritance of semantic content from the content of the percept.

The principal reason I favor this account over the unmediated one is the one mentioned briefly above: I don’t think it makes sense to think of the mind actively selecting an external object. We saw above that the basic semantic connection between mind and world can involve a MWM or a WMM. Selection is a MWM, involving an intentional mental act. By virtue of this fact, the objects upon which it operates must themselves be mental entities, mental representations. Put another way, when the cognitive subject selects something, she/it must know what she’s selecting, which entails that what is selected is represented. The immediate relation, then, obtains between the subject and the representation, and by virtue of that relation and the semantic content of the representation, the subject counts as selecting, attending to, or demonstrating the object of that representation.

Lest one think this sets up a vicious regress, note that the basic semantic relation that connects the mind to the world need not involve an MWM; and, if I’m right, it cannot. Representation itself is not a mental act or process, and therefore need not relate the mind to another representation. In fact it is crucial that one not treat representation as a mental act. Thirty to forty years ago it was standard for critics of the computational theory of mind to object that the model involved a regress: if the mind is representing, then there has to be a subject for whom it is representing, and then there has to be a process of understanding or interpreting, which, being a mental process itself, entails new representations, and so on. The response was to point out that the mental representations are used, manipulated, processed, transformed, etc. - but not themselves interpreted. To represent x in this sense is just for the mind to token a representation that has x as its content. No process or act of interpretation or representation is needed.

Notice, however, that sometimes we do represent as an explicit act or process. For instance, at the Passover Seder we set out various objects and claim that they represent certain significant objects and events from ancient history. So, for instance, we say things like “this bitter herb represents the misery the Israelites suffered as slaves in Egypt.” Here I am performing a mental act of using this herb to represent suffering. In this case, however, it seems clear that in order to perform the intentional act of using the herb to represent suffering, I need to already possess representations (in the first, basic sense) of both the herb and the suffering. My mental act is then one of associating the two objects/events in the world by way of acting on my representations of them. How could I intend to use the herb to represent suffering if I don’t have an independent way of representing both the herb and the suffering? What would it even mean to, as it were, directly associate the two without mediation by the requisite mental representations? I claim the same requirement of prior representation holds for selecting/attending/demonstrating as for this non-basic sense of representing.

If the foregoing is right, then what grounds mental singular reference must be an RMM that is also a WMM. While the mind cannot “grab” objects in the world directly - that is, without first representing them - objects in the world can causally “grab” mental representations and thereby constitute their semantic contents. In our case of the md [that fly], we see that mechanism at work in the relation between the fly itself and the visual percept of the fly. Here we find the ground-level connection between mind and world that serves as the foundation on which other mental representation is built.

Before ending this section and turning to a consideration of type-demonstratives, I want to briefly describe at least one theory, Zenon Pylyshyn’s, of how percepts acquire their referents. While Pylyshyn’s view is certainly controversial within the psychological community, it does illustrate some of the themes I’ve been discussing and provides a nice model of how mental demonstration and perception might work together to secure singular reference in thought.

Pylyshyn (2001 and forthcoming) argues that already in early vision distal objects are
represented, and that they 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”.

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 a cue, in the form of 3 or 4 of the squares lighting up, appears, 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, in view of the considerations developed above, these metaphors pull in opposite directions. On the one hand, he often speaks of an object “grabbing” a FINST, which is where I borrowed the term from in my discussion above. The idea is that an object presents a certain collection of spatio-temporally distributed features to the visual system, and when this happens, 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 qualitative information is represented within the FINST, though it plays a causal 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 pure, unmediated, RMM-WMM. This model of the object “grabbing” the visual percept is the one I want to tentatively adopt for my account of the percept-object relation that ultimately grounds the md-object relation.

On the other hand, Pylyshyn also compares FINSTs to demonstratives, and talks of FINSTs as mechanisms for selecting objects. If what I’ve argued above is correct, then this can’t be the right way to picture the relation between FINSTs and objects. In fact, I think Pylyshyn’s position is best captured by the world-to-mind “grabbing” metaphor presented above, and his remarks about direct selection of objects, as well as his comparisons between FINSTs and demonstratives, can be explained in a way that is not really inconsistent with my position.

When it comes to the comparison to demonstratives, I think what’s essential to the comparison is that FINSTs, like demonstratives, do not contain either descriptive semantic content or descriptive metasemantic content - that is, they are semantically directly referential and they possess RMM’s, not AMM’s. In this respect, I too agree that FINSTs resemble demonstratives. It’s just that I want to insist that the RMM’s by which md’s attach to their referents are mediated through visual percepts, or 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 (1980) “feature-integration” theory. On the latter theory, 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 vision is attending to, selecting, an object, it 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.

If Pylyshyn is right, then the complete model of token-demonstrative thought is as follows. Through its causal influence upon the early visual system, particularly by stimulating the right combination of early feature detectors, an object causes a visual tag, a FINST, to be formed. By thus “grabbing” the FINST, the object becomes its semantic content. This is an example of a WMM and an unmediated RMM. Once the FINST is formed, it then continues to represent that object through changes in its features, and serves as a singular term on which to hang certain predicative, featural contents. Unlike higher-level singular representations, however, this referential link endures only for limited periods of perceptual contact with the object.

While these visual tags, the FINSTs, endure they are available as objects of visual attention. When one such tag is selected, its content then becomes available for inclusion in a demonstrative thought. The thought contains a placeholder for an attentive pointer to the visual tag, and through this arrangement the thought is about the object represented by the tag. The md, which just is this pointer, has the distal object as its semantic content. The metasemantic mechanism by virtue of which it acquires this content is an MWM, a mediated RMM. I think about the fly by visually attending to it and incorporating what I’m attending to into my thought. But the way I grab ahold of the fly is by pointing at a visual representation that was already grabbed by the fly. Thought meets the world in the middle, at the percept.

One final point before turning to type-demonstratives. I’ve gone into detail about how my view of demonstrative thought coheres with Pylyshyn’s view about visual object representation for two reasons: first, I’m attracted to his theory, and so want my view to be consistent with it, and second, it very neatly fills in the gap in the metasemantic story I have to tell about md’s. If they have mediated RMM’s, one needs an account of the mediation. However, if it turns out that the right story in the end is that one visually attends to spatial regions after all, the main outlines of my view about demonstrative reference in thought will not be seriously affected.

4. 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 say 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 the other thought, serves to demonstrate a type of color, not any token paint.

At the beginning of the paper I mentioned that just as we can assume there are mental thought-constituents corresponding to token-demonstrative expressions like “that fly”, so too we can assume there are mental thought-constituents corresponding to type-demonstrative expressions, like “that color”. The question is, can we treat mental type-demonstratives similar to the way we treated their token cousins? In particular, does the pointer analogy make sense in this case?

In fact I think it does, and it helps to explain 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 the 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

10 11. A lot will depend on just how object representation in vision is built up from the initial representation of features instantiated in spatial regions. The relation between demonstration and visual attention could also become more complicated. This is why I like Pylyshyn’s story, because in many ways it’s much simpler. However, my main point concerning the limits of MWM’s, that they operate on mental representations, and that the ultimate ground of mental reference must involve a WMM, would still stand.
of the house by looking at the house in the dark, not seeing its color, and still thinking “that color”, 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. But notice that if we analyzed the type-demonstrative as essentially equivalent to having a pointer to the percept of the house and then thinking “the color of that”, we’d have no way of distinguishing between these two ways of thinking of the color. But can one point directly at the color, as a type? Doesn’t it have to be a matter of pointing at the individual and then thinking of the color of that individual?

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. However, 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 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 md points not at the FINST, but at the relevant feature-predicate, the embodiment of one’s perceptual experience of the feature in question. The md, though pointing at a predicate-like representation, itself still plays the role of a singular term in thought. So the result is a thought about a demonstrated property.\(^{11}\)

So the difference between the two ways of referring to the house color turns out to be quite dramatic. The first way, which involves seeing the color and thinking [that color], is realized through a pointer-mechanism that addresses the perceptual representation of the color instantiated in an object (or in a certain spatio-temporal region). I will call this a “true type-demonstrative.” The second way, thinking [the color of that house] when seeing the house in the dark, is a definite description a constituent of which is a token-demonstrative to the house. I’ll call this sort of demonstration a “pseudo type-demonstrative.” Both token-and type-demonstratives are pointers, one to percepts of objects - FINSTs perhaps - and the other to perceptual representations of features. In this sense, as one looks at the color and thinks [that color] one has it in mind in as direct (and also as indirect) a way as one has the fly in mind as one sees it on the wall and thinks [that fly]. To the extent that pointing in this specific sense is a mechanistic realization of acquaintance, it shows how, as Russell believed, we can be acquainted with universals as well as particulars.

5. 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 type-demonstratives cannot serve the function of concept acquisition/formation. I’ll begin by briefly outlining the three cases.

First, consider how McDowell (1996) employs type-demonstratives. McDowell wants to argue that

\(^{11}\) 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.
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 vis a vis 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. 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 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 their respective 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.

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

---

12 Perry (2001) also makes a similar move, though he doesn’t explicitly address the explanatory gap, but rather the conceivableability argument. However, his response to that argument could be used to address the gap as well.
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. According to the account in section 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.

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.

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 when they characterize phenomenal concepts as involving type-demonstratives are true 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 founds on the same obstacle. One must already be aware of, and therefore 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. 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, pseudo-type-demonstratives are clearly not 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 it’s reference is determined along the lines I outlined in section 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 in this paper is that demonstration/selection/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. Pointing, demonstrating, as I’ve characterized it, is a mental operation, and by virtue of that fact it takes 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.¹³ 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.

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 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 the one Jerry Fodor (unpublished) has pointed out already, and it’s one we’ve seen already when discussing using pseudo-type-demonstratives to form phenomenal concepts. 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

¹³ See Levine (2007) for more general discussion of this strategy.
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.

REFERENCES


