Concepts as Rules: A Kantian Proposal

Hannah Ginsborg

VERY ROUGH DRAFT: PLEASE DO NOT QUOTE OR CIRCULATE WITHOUT PERMISSION

In an overview of the topic of concepts (2005), Stephen Laurence and Eric Margolis distinguish three possible positions that might be taken regarding the ontology of concepts. Concepts could be mental representations, they could be abstract objects (in particular Fregean senses), or they could be abilities, in particular abilities for sorting or inference. Laurence and Margolis favour the view that concepts are mental representations, or what they call the psychological view. The main advantage of this view, as they see it, is its superior explanatory power: it accounts both for the productivity of thought, that is for the way in which human beings are capable of entertaining an infinite variety of thoughts, and for the fact that rational mental processes can be realized in a physical system.[1] This last point in particular represents a contrast with what they call the semantic view, on which concepts are abstract objects, and more specifically Fregean senses, that is, constituents of Fregean thoughts. The trouble with the Fregean sense view, as they see it, is the obscurity of the notion of grasping a Fregean thought or its constituents. The relation of grasping cannot be causal, since senses fall outside the realm of physical causes and effects. But if it is not causal then it is, they say, utterly mysterious (2007, 580). The third view, namely the view that concepts are abilities, is not discussed by them in detail. But they suggest that it is largely motivated by mistrust of the idea of mental representations, and also that it cannot account for the productivity of thought. So they see the case for the mental representation view, or the psychological view, as overwhelming.

Are the three positions identified by Laurence and Margolis in genuine disagreement? Some philosophers think they are not. Christopher Peacocke, for example, whose book A Study of Concepts assumes an understanding of concepts as with Fregean senses, allows that there could be another sense of

[1] Their reasons for favouring this view are given in a number of papers, in particular their 1999 and 2007, on which I will be drawing in what follows.

the term "concept" on which concepts are mental representations.[2] Moreover, it is common both for philosophers who identify concepts with Fregean senses, and for philosophers who identify them with mental representations, to speak of concepts, in addition, as though they were abilities or capacities. Peacocke, again, speaks of concepts as things that can be mastered or possessed, a way of speaking which is more apt for abilities than for Fregean senses or other abstract objects; and indeed he is treated by Jerry Fodor (2004) as a prime example of commitment to what Fodor calls concept pragmatism, the view that concepts are capacities for sorting or inferring. But Fodor, the prime exponent of the mental representation view, also sometimes describes concepts as a kind of ability: possessing the concept dog, for example, is having the ability to think about dogs as such.[3] Leaving aside the question of whether concepts are abilities, though, Laurence and Margolis, like Fodor, take the dispute between the mental representation or psychological view, and the Fregean sense or semantic view, to be a genuine one. And Laurence and Margolis warn that there are dangers in being too quick to see an apparently substantive dispute as terminological: such an approach across the board might lead us to treat disputes about goodness or free will as terminological also (2007, 589n10).

If there is a genuine dispute here, then we must be able to identify a core of pretheoretical intuitions associated with the notion of a concept: it must be that there are some constraints, implicit in our use of the term "concept," which a theory of the ontology of concepts has to do justice to. As Peacocke has pointed out, ordinary non-philosophical usage does not seem to help us here.[4] The term concept, he says, is a "philosophical term of art." Still, that does not mean that there cannot be a pretheoretical notion of what a concept is which applies within philosophical and (for that matter psychological) tradition. This would be pretheoretical in the sense that it would precede specific ontological views like the ones outlined by Laurence and Margolis, and such views would be answerable to it. Laurence and Margolis indeed provide such a pretheoretical notion: they suggest, in their encyclopedia article, that concepts, pretheoretically understood, are "the constituents of thoughts." But this does not get us much further, since how one understands this suggestion depends on what pretheoretical content one attaches to the term "thought": in

[2] 1992, 3. Alex Byrne also takes this view (2005, 231).
[3] Although for Fodor abilities do not constitute having concepts "in any interesting sense" (1998, 3).
[4] As an illustration he quotes the line from Annie Hall which Woody Allen gives to a character in the entertainment industry: "Right now it's only a notion, but I think I can get money to make it into a concept, and later turn it into an idea" (1992, 1).

particular it depends on whether one thinks of thoughts as entities or occurrences in individual minds, or more along the lines of Fregean thoughts, that is, as abstract entities. It's clear that Laurence and Margolis take the former line, and that this naturally supports the mental representation view. If on the other hand one thinks of thoughts as Fregean, then concepts will come out as Fregean senses. [The question about the right way to understand thoughts and thinking does, I think, reflect a substantive philosophical dispute. But it is not obvious that the dispute is about the nature of concepts as such. Rather it would seem to be about the extent to which thinking can be understood in naturalistic terms, as a purely psychological process, as opposed to something which requires a grasp of propositions and meanings which are not naturalistically conceived.]

Is it possible to be more specific about the pretheoretical sense of the term "concept," in a way which gives some content to the dispute about ontology? One way to get clearer about the pretheoretical intuitions associated with the notion of a concept is to look at how the term "concept" has been used in the philosophical tradition. And here the obvious place to begin is with the philosopher who is, I believe, largely responsible for introducing the term "concept" into recent philosophical discourse, namely Kant. The meaning of this term for Kant can be seen most clearly in the context of the contrast he draws between concepts and intuitions. Intuitions are singular representations which relate immediately to their objects: they are the representations through which we come to stand in cognitive relations to particulars, specifically to particular spatio-temporal objects. Concepts on the other hand are universal or general representations; they do refer to particular objects, but only in a mediated way, in so far as they represent objects as having something in common (Logic §1, KdrV A320/B377). So concepts -- and this is the first of two points I want to emphasize about Kant's view -- allow us to represent individual things as having general features; as I'll put it for short, they make possible the representation of generality. Kant goes on, and this is a second, although I believe philosophically related, point, to make a connection between concepts and judgments, which he puts by saying that concepts are "predicates of possible judgments" (KdrV A69/B94). For Kant the idea of judgment is tied to the idea of a claim to universal intersubjective validity. When you make a judgment, you eo ipso claim to be judging as you ought, in a sense that implies everyone ought to judge the same way that you do: it is part of your act of judging that you take your judgment to demand universal agreement. So the idea of concepts as predicates of possible judgments is

the idea that in representing something conceptually you are in the first instance making a claim that everyone else ought to represent it in the same way you do. This doesn't mean that all conceptual representation involves judgment -- it is possible simply to entertain a conceptual content without judging that things are as they are represented to be. But the idea of judgment has priority: it is in virtue of their role in judgment that concepts make possible the more neutral entertaining of thoughts, or the adoption of non-judgmental attitudes to them such as desire or fear.

Kant's notion of concept is more specific than the notion of a "constituent of thoughts" in ways corresponding to both of these points. First, a concept is specifically predicative: it is the kind of thing which can be expressed only by a general term. By contrast, constituents of thoughts as understood by Laurence and Margolis include mental representations corresponding to singular expressions like "Fido" or "George Orwell". Accordingly, as they see it, a concept is any subsentential mental representation, whether or not it plays a predicative role. The only reason, they suggest, not to identify concepts with representations corresponding to whole sentences is that sentences are too big (1999). Second, for Kant we can understand what a concept is only in terms of its role in judging. By contrast, the idea that concepts are constituents of thoughts doesn't on its own imply any special priority for belief or judgment over other attitudes to thought-contents. Concepts figure in the contents of propositional attitudes, which include judgments, but there is no privileged relation between conceptual representations and judgments as opposed to propositional attitudes of other kinds.

Now my point in introducing Kant's view of concepts has been to try to clarify some of the intuitions that are associated, in the philosophical tradition, with the use of the word "concept." I am not saying that we should simply conclude, from Kant's use of the term, that this is what the term "concept" now means, and that any present-day theory of concepts has to be answerable to it. On the other hand, I think there's a case to be made for saying that the two Kantian considerations I've mentioned -- the idea that concepts enable us to represent things as having general features, and the idea of an intimate connection between conceptualization and judgment -- correspond to at least one important understanding of the term "concept" which is still presupposed in philosophical discussion. Part of the case is historical. Kant's use of the term "concept" was highly influential in the subsequent tradition, in particular via Frege, whose own notion of concept retains both the connection with the representation of generality -- through the idea that a

concept is the meaning or Bedeutung of a general term or predicative expression -- and the connection with judgment. So it is hard to suppose that subsequent uses of the term "concept" in the post-Fregean philosophical tradition were free of those implications. [Admittedly the identification of concepts with Fregean senses does signal a departure from Frege's own notion of a concept. This is first because for Frege himself concepts belong at the level of reference rather than at the level of sense, and second because the notion of a Fregean sense applies indifferently to singular terms and to predicative expressions, whereas for Frege himself concepts are tied exclusively to predication. Nonetheless, I do not see the first point as undermining the connection between concepts and judgment, and I do not see the second point as representing an abandonment of the idea that the central use of concepts is in representing things as things as members of general types or kinds.] And it seems to me that much contemporary philosophical discussion which is not specifically about concepts, but which still makes use of a pretheoretical notion of "concept," presupposes an understanding of concepts both as representing generality, and as tied to judgment or belief. This is particularly clear in discussions of whether perception has nonconceptual content. When Evans, Dretske, Peacocke, and others defend the claim that the content of perception is either wholly or partially nonconceptual, they take themselves to be arguing that it is importantly unlike the content of beliefs or judgments, and this suggests that their notion of a concept remains linked to the Kantian tradition.

In this paper I am going to appeal to the pretheoretical notion of a concept which I have drawn from Kant to try to shed some light on the debate about the ontology of concepts. Part of what I'm going to try to do -- and this will be the topic of the next section -- is to argue that the mental representation view fails to do justice to the two Kantian conditions I've mentioned on something's being a concept. So if we want to retain an understanding of the term "concept" which is continuous with its original sense in Kant, then**,** I'll argue, concepts cannot simply be identified with mental representations. But I do not want to endorse the Fregean sense view either, or, to be more specific, I do not want to endorse it if Fregean senses are understood as entities in a kind of Platonic "third realm."[5] This is because I am sympathetic to the worry that Margolis and Laurence raise about how we can grasp such senses: namely, the worry that the kind of cognitive relation involved in grasping such senses is mysterious. So I will propose, in the third

[5] There is, I think, another way of understanding Fregean senses which brings them closer to Kantian concepts as I understand them [cf. work by Tom Ricketts], but I will not try to go into this here.

and final section, an alternative view of concepts, which is itself Kantian: namely that a concept is a kind of rule. Kant himself identifies concepts with rules for what he calls the synthesis of the imagination. This view is hard to explain briefly, but in the case of an observational concept like dog the concept can be identified with a rule for forming perceptual images. I am not going to commit myself to this particular aspect of Kant's view, and nor am I going to make a claim about what, in general, concepts are rules for: that is, I won't try to give a general account, applicable to all concepts, of what activity it is for which concepts serve as rules. But I am going to propose a more specific claim regarding the special case of observational concepts -- by which I mean those concepts which can by and large be applied on the basis of perceptual experience, such as dog, green, platypus, or clock. A concept of this kind, I shall suggest, can be identified with a rule for discrimination, that is sorting. You possess the concept dog if you grasp a rule for sorting dogs from non-dogs. Now of course the claim that a concept is a rule, whether for sorting or for anything else, doesn't on its own constitute much of an advance on the ontological issue, since we can raise exactly the same question about the ontology of rules as we can about concepts: are they mental representations, abstract objects, or abilities? And if they are abstract objects, as indeed I will suggest they are, then we face the same difficulty that we face for Fregean senses, that is, of how to understand what it is to grasp them. But I am going to be drawing on the idea of normativity which is implicit in the notion of a rule to try to provide a satisfying answer. Roughly, I am going to say that you grasp a rule for sorting dogs from non-dogs if, and only if, you can discrminate dogs from non-dogs with the awareness that you are discriminating appropriately.

II

In this section I want to argue that the mental representation view fails to do justice to the two Kantian desiderata I mentioned: that concepts enable us to represent generality, and that to represent something conceptually is paradigmatically to make a judgment about it, in the sense of a claim which demands universal agreement. First, though, I'll offer a brief sketch of the mental representation view, at least the version of it which I think Margolis and Laurence have in mind when they include it in their taxonomy of ontological theories. The view derives from Fodor's "language of thought" hypothesis, on

which thinking, construed broadly as the entertaining of propositional attitudes, is to be explained in terms of a language-like system of internal mental symbols which is referred to as the language of thought, or Mentalese. A propositional attitude like a belief or a desire consists in a relation to a sentence of Mentalese: such sentences consist, in a way analogous to sentences in a public language like English, of strings of internal symbols which are analogous to words. Concepts are, in the first instance, to be identified with these word-like symbols, and with subsentential combinations of them. So when you come to believe, for example, that Fido is a dog with black spots, you come to stand in a certain relation to a Mentalese sentence meaning "Fido is a dog with black spots." The concepts involved here are the symbols meaning Fido, dog, black and spot, and also subsentential combinations of them, such as the symbol for dog with black spots. Believing, as opposed to desiring, that Fido is a dog with black spots is a matter of the sentence's being located in your "belief box," which is a metaphor for its tending to have certain effects on you rather than others. For example, having the sentence in your belief box amounts to something like this: if you like black-spotted dogs, then you might pet Fido, or if you are afraid of black-spotted dogs, then you might run away from him. Having it in your desire box, on the other hand, means that you might take Fido to the dog parlor to have black spots painted on.

Now concepts on this view are, in the first instance, particulars. This means, among other things, that they are located in individual minds, which is a point which very clearly distinguishes this view from the view that concepts are Fregean senses. But like words or other expressions in a language they can also be understood as types. What there is in your mind, then, when you have a belief involving the concept dog, is what we might think of as a particular inscription or token of the Mentalese word DOG. We can have different tokens of that same Mentalese word, just as the English word "dog" can be tokened on different occasions. Moreover, the token words of Mentalese, like those of English, can be typed either nonsemantically, for example through formal features analogous to spelling, or semantically in terms of their meaning. The two tokenings of DOG which occur in your mind when you think, first, "Fido is a dog" and then "Rover is a dog" are of the same formal type -- they're both, as it were, spelled the same way. But they are also of the same semantic type because they both have the same meaning or content, namely dog. The view that concepts are mental representations is compatible with a range of different accounts of how mental representations get their semantic properties, and thus of why it is that any token of the syntactic

type DOG means dog rather than, say, grapefruit or oscillator. But on Fodor's version of the view, with which I take Laurence and Margolis to be broadly in sympathy, it is through the holding of lawlike correlations between the instantiation of a property in the world and the tokening of a concept in the mind. Very roughly, the mentalese word DOG means, or has as its content, dog, because it is a law that dogs cause tokenings of DOG. That's what allows us to describe tokens of the mentalese word DOG as tokens of the concept dog. As Fodor puts it: "to a zeroth approximation, the fact that DOG means dog is constituted by a nomic connection between two properties of dogs; viz being dogs and being causes of actual and possible DOG tokenings in us" (1998, 73). That rough view is refined by Fodor to account for the fact that other things, such as disguised cats or thoughts about kennels can also cause DOG to be tokened, without DOG tokens coming to mean anything other than dog. But the details of how the view gets refined, in particular through Fodor's asymmetric dependency thesis, don't matter for our purposes. [6]

Does this picture do justice to the idea that a concept is what enables me to represent a particular as having a general feature, so that, for example, in tokening the concept "DOG" I am thinking of Fido as a dog? One might initially be suspicious that it doesn't, on the ground that concepts, on Fodor's view, are supposed to be particulars, yet it is hard to see how something particular could amount to the representation of a general feature. Although Fodor and other defenders of the mental representation view are not explicit about this, I assume that it is a consequence of their view that if you think on one occasion that Fido is a dog and on another occasion that Rover is a dog, your thinking on these two occasions involves two different token concepts. This seems to follow from the analogy with language: if you say or write the words "Fido is a dog" and then the words "Rover is a dog" the two occurrences of the word dog are different linguistic tokens. So it might be wondered how on each occasion we can be thinking the very same thing about Fido and about Rover, if each thought involves a different token. This particular worry is related to a worry which is explicitly considered both by Fodor, and by Margolis and Laurence, the worry about how the same concept can in principle be entertained by two different people -- how concepts can be public. Fodor takes it to be a pretheoretical demand on a theory of concepts that it make sense of concepts'

[6] Here is one formulation of the asymmetric dependence view: DOG means dog if (a) it is a law that the property dog causes DOG tokenings and (b) if other causal relations between properties and DOG tokenings asymmetrically depend on this law, where X's causing Ys "asymmetrically depends" on a law L iff X's causing Y wouldn't hold but that L does, but not vice versa. (The formulation is adapted from Loewer and Rey 1991 [add ref.], who introduce the shorthand "nomological locking" for the relation between DOG and dog.)

being public, that is, of the possibility of our sharing concepts (1998, 28-29). But two different people cannot share the same token concept.

Now Fodor's reply to the worry about publicity is that two different people can entertain the same concept in the sense of concept type. The same reply can presumably be used in the case of the present worry about generality: I am thinking the same about Fido as I am about Rover because my DOG tokens are tokens of the same concept type. So there is indeed generality somewhere in the picture: my particular concept tokens are instances of a general concept type, so that my thought about Fido and my thought about Rover are the same type of thought. I am thinking the same about Fido as I am about Rover in the sense that I am instancing, in each case, the same way of thinking: my thoughts are both instances of the same general kind. But this on its own doesn't explain how I am representing Fido and Rover themselves as belonging to the same kind.[7] For me to represent Fido and Rover as having the same general property it is not enough for my representations to have the same general property. Rather, it would seem, the general property I represent Fido and Rover as sharing must somehow figure within the intentional content of each of my representations, rather than being a property of the representations themselves.

The mental representation theorist will, I think, accept this point. He or she will allow that appeal to the possibility of typing representations isn't on its own sufficient to address the concern; rather we have to appeal to the part of the account which specifies the semantics of the representations. That is, we will have to appeal to semantic, not just to syntactic typing. In the version of the account that we're considering, that would mean appeal to the nomological correlation between dogs in the world and tokenings of DOG in people's minds, {or more precisely, between the property of being a dog and the property of being a DOG token}. When I token DOG in connection with Fido, I am representing Fido as a dog because my token belongs to a syntactic type which is nomologically correlated with the general property of doghood. When I token it in connection with Rover, I am representing Rover as a dog for the very same reason. So now we have another kind of generality in the picture: the generality of the property of doghood with which my DOG tokenings are nomologically correlated. And we can also say that it is a general law that, when I am in the presence of doghood, I produce a DOG token.

[7] As Kent Bach appears to suggest (2000).

But I still don't think that this gets us all the way to explaining how it is that I come to represent Fido or Rover as instantiating doghood. I think that the generality is, again, in the wrong place. For there is, on the face of it, a gap between having a general lawlike disposition to respond discriminatively to the presence of dogs, and being able to represent something as being a dog. Mere sensitivity to doghood isn't, again on the face of it, sufficient for the capacity to ascribe doghood, or indeed any other property possessed by dogs: you can be a reliable detector of dogs without being able to represent the things whose doghood you detect as dogs, or indeed as anything else at all.[8] And I don't see that this version of the account has the resources to show how my tokening of DOG, when I see Fido, is anything more than a mere detection of Fido's doghood. Fodor, I think, wants to say that somehow, through the nomological correlation between DOG tokenings and doghood, my DOG tokens come to have intentional content; they come to represent the property of doghood. That nomological or counterfactual correlation is supposed to make it the case that, each time a particular thing actually brings about a tokening of DOG, I represent the thing as a dog. But it's not clear how the fact that, if I were to be in the presence of a dog, I would token DOG, makes it the case that when I do token DOG in the presence of some object, I am thereby representing it as being a dog. Rather, it seems, I am simply producing a response of a kind which I typically produce to dogs, and hence -- if the thing I'm responding to is in fact a dog -- registering its doghood. If Fido isn't a dog -- if he is a disguised armadillo, say -- then I haven't misrepresented him either: rather, the mechanism which makes me sensitive to doghood has just led me to produce a doghood-indicating response to something which isn't a dog.

It is of course open to the defender of the mental representation view to reject the distinction on which that last point relies, that is, the distinction between responding discriminatively to doghood and representing something as a dog. In fact, one might say that the whole point of informational semantics is to deny that distinction. For the informationalist, representing something as a dog just is having a

[8] My worry here should be distinguished from the kind of worry which is often raised about views of this kind, namely that they don't account for different modes of presentation with the same extension, e.g. Fido's being a dog as opposed to a barker. My worry is that they don't account for the possibility of representing Fido as anything at all, not that they don't distinguish different but coextensive ways of representing Fido. The contrast I have in mind here is emphasized both by Robert Brandom (in his distinction, for example, between sentience and sapience, and in the contrasts he repeatedly draws between, say parrots, thermostats and pressure plates on the one hand, and human beings on the other) and by John McDowell. But, in part for reasons suggested by McDowell, I do not think that Brandom's inferential approach is any more successful than Fodor's in doing justice to the contrast. (This is a point which requires much more discussion than is possible here.)

representation of a formal or syntactic type which is nomologically correlated with doghood: the nomological correlation is precisely what does the work of providing representations of that type with semantic content. But part of what I am trying to do here is to point out the costs associated with making this kind of move, one of these costs being the abandonment of the kind of generality which is associated with the notion of a concept in the Kantian tradition. And it seems to me also that, in denying the distinction between, so to speak, detections of doghood and ascriptions of doghood, the mental representation theorist also has to abandon another intuitively plausible distinction, which we can put very roughly as the distinction between responding psychologically to something, and making a judgment about it. This distinction corresponds to the second of the two Kantian requirements on something's being a concept. According to that requirement, to represent something conceptually is in the first instance to make a claim about it: in representing Fido as a dog I am by default taking or judging Fido to be a dog, and thus implicitly making a claim on the agreement of others. (While it is of course possible for me to withhold that claim, as when I merely entertain the hypothesis that he is a dog, the possibility of my doing that is parasitic on the possibility of my using the concept in a way which, as Kant puts it, demands universal agreement.) Conceptual representation construed in this way requires more than "mere responsiveness" to Fido's real or apparent doghood even if the responsiveness is mental rather than behavioural. But the mental representation view, at least on the informationalist version we are considering, does not on the face of it respect that Kantian requirement. That I token a mental representation of a kind which is keyed to the presence of doghood doesn't in itself imply my taking it that I or anyone else ought to token that representation, hence it doesn't imply that I demand agreement for my tokening. Nor, as far as I can tell, does that implication follow from my tokening it, more specifically, in my belief box. That the sentence FIDO IS A DOG is tokened in my belief box has implications for how I will act and for what other sentences I will token, but it doesn't so to speak endow my tokening with a claim to universal agreement.

This worry can be illustrated in connection with Margolis' and Laurence's handling of the question about the publicity of concepts, that is the question of how two people can share the same concept. As we saw, they take it to be sufficient for publicity that the two people token the same type of concept. In developing this point, they say that it is not necessary for communication that each person know what concept is associated with each of the other person's words. All that is needed, they say, is coordination of

words with concepts: it is enough, as they put it, that our "mechanisms of language acquisition... reliably attach mental representations with the same content to the same words," so that there are "reliable correspondences between language and thought" (2007, 589n11). But this doesn't address a more fundamental Fregean requirement on communication, which I take to be derived from Kant's distinction between mere psychological responsiveness and judgment: namely, that communication allow for the possibility of genuine agreement and disagreement. For a Fregean, communication doesn't just require that there be sameness or difference in the psychological states we express under the same external circumstances, it also requires that we be able to express claims about those circumstances which can accord or conflict. It seems to me that the mental representation view fails to capture this requirement; it cannot do justice to the idea that if, on seeing Fido, you token DOG and I token CAT, we are making genuinely competing claims. So the Fregean demand on communication is not satisfied. To put the point vividly if crudely, our situation is like that of two parrots, trained to say "dog" and "cat" to dogs and cats respectively, who respond differently to a given animal of indeterminate appearance. When the first parrot says "dog" and the second parrot says "cat" they are presumably also expressing differing psychological responses to the animal presented to them, psychological responses which may well consist in the tokening of mental representations like the Mentalese DOG and CAT. However, at least on the intuition underlying Kant's view of concepts, they don't count as disagreeing with each other; rather, they are merely responding differently. And by the same token -- to go back to the first condition on conceptual representation -- they are not representing the animal as having any feature, whether doghood, cathood or anything else.

III

I argued in the previous section that the mental representation view of concepts, at least on the version we were considering, doesn't respect the Kantian constraints on the notion of a concept. The view that concepts are Fregean senses does seem to respect those constraints, but as we saw at the beginning, it is open to criticism on the grounds of its lack of intelligibility: if concepts are entities in some kind of third realm, then it seems mysterious that human beings should be capable of grasping them. Can we do justice to the Kantian requirements in a way which avoids this criticism? I want to suggest that we can. Let's

begin by supposing, as on Fodor's version of the mental representation view, that representing Fido as a dog does indeed involve tokening a mental representation, DOG, which is nomologically correlated with the property of doghood. I argued in the previous section that that does not amount to representing Fido as a dog. If Fido is a dog, then in responding with DOG we are detecting his doghood, but that does not amount to ascribing or predicating doghood of him. That might seem to commit us to something like the Fregean-sense view. In order for our tokening of DOG to represent Fido as a dog, it might seem, DOG must mean dog independently of its nomological correlation to dogs: that is, it must have a meaning which is somehow fixed independently of how we actually token it.

But I want to suggest another possibility. My suggestion is that you are in a position to represent Fido as a dog by tokening DOG insofar as your DOG tokenings are not only nomologically correlated with the presence of doghood, but also involve a normative element: they involve your taking yourself, in tokening DOG, to be responding appropriately to your circumstances. DOG thus comes to mean dog in virtue of the fact that the presence of dogs reliably causes you not only to token DOG but do so with a consciousness of normativity: a consciousness of doing as you ought given both the presence of Fido and the history of your previous tokenings of DOG. I want to suggest that it's this consciousness of normativity in your tokenings which makes the difference between merely responding to Fido's being a dog, and representing him as being a dog. For in tokening DOG with the consciousness of normativity you are in effect taking there to be a normative fit between Fido and the tokening of your mental representation: you are taking Fido to make that representation appropriate. And in so doing, you are taking Fido to have something in common with all of the other objects with respect to which you do, or would, take the same representation to be appropriate. You are, so to speak, taking Fido to be the sort of thing with respect to which DOG ought to be tokened, given your previous tokenings of DOG. This means that you are satisfying both of the two Kantian constraints on conceptual representation. You are representing Fido as having a general feature in common with the other things to which you have responded by tokening DOG. And in tokening DOG you are in effect making a claim on others' agreement, that is, a judgment: you are claiming, in tokening DOG, that anyone else in your situation ought to token DOG as well.

I have presented the proposed view as a modification of the mental representation view, but it can also be thought as a modification of the view that a concept is a kind of discriminative capacity. Seen in

this way, my suggestion is that to grasp a concept -- at least the kind of ostensively acquired concept which we are considering -- is to have a capacity to produce some type of discriminative response to a kind of thing, where each response you produce involves a consciousness of its own appropriateness given the circumstances in which it is produced. We could imagine a simple version of the discriminative capacity view on which grasping the concept dog was a matter of being able to respond to dogs by saying some word, for example the word "dog." On this view, a parrot who had been trained to respond to dogs by saying "dog" would count as possessing the concept dog. I am suggesting that while this capacity is not sufficient for concept-possession in the sense of "concept" specified by the Kantian requirements, it is sufficient if we give our characterization of the capacity what I've called elsewhere a "normative twist": that is, if we add to our characterization the proviso that the subject takes each utterance of "dog" to be appropriate given her circumstances. More specifically -- to unpack what I mean by "her circumstances" -- the proviso requires that she take it to be appropriate in the light of the presence of the particular object, say Fido, and in the light of her previous usage of the word "dog." It is not essential to the view that the discriminative response be the utterance of a word in a public language: it could, as in my initial presentation of the view, be the tokening of a symbol of Mentalese, or it could be some other psychological or behavioural response. All that is essential is that the response be reliably correlated with the presence of the property, and that, in producing it, you are not just, so to speak, blindly reacting to the object, but rather producing a response with a sense of its fitting the object in the light of your previous responses of that type.

Now I shall suggest in a moment that the view I have characterized amounts to the identification of concepts with rules.[9] For, I'm going to suggest, having the capacity to respond discriminatively to dogs with the awareness that one's response is appropriate in the light of one's previous responses just is grasping

[9] Note that Brandom also identifies concepts with rules, and indeed takes himself to be following Kant in so doing. See, for example, the identification of concepts as norms determining the correctness of inferential moves (2000, 29) and his characterization of Kant (2000, 80) as identifying concepts as the rules that determine what knowers and agents are responsible for. It is not possible here to go into the differences between Brandom's view and the view I am proposing. But a starting-point is to note that for Brandom the activity which is governed by the rules is itself an activity of applying rules: relatedly, it is an activity of making judgments which are already construed as conceptual, and indeed as involving the very concepts which, qua rules, govern the activity of making the judgments. On the view I am proposing, by contrast, the activity can be understood in a way which does not require appeal to the rules or concepts which govern it. A related point of difference is my commitment to the idea of primitive normativity, which I try to sketch in the text following: I take it that Brandom would not endorse the idea that we can make out a sense of "ought" which does not depend on grasp (even "implicit" grasp) of a rule.

a rule for discriminating dogs. But to introduce that suggestion I need first to consider what I take to be the most obvious objection to the view as I have so far articulated it. It goes like this. How can you take your response to Fido to be appropriate, that is, how can you take yourself to be responding to Fido as you ought, if you don't first have in mind a rule to which you take yourself to be conforming? More specifically, how can you take your saying the word "dog", or tokening the mental representation DOG, to be appropriate to your circumstances if you don't antecedently have in mind a rule which dictates that you ought to apply the word "dog", or the mental representation DOG, to dogs? And if you do have to have such a rule in mind, then it would seem that you must already grasp the concept dog. For the rule must tell you that you are to apply DOG to dogs, and without a grasp of the concept dog it would seem that you are not in a position to understand the rule. The upshot is that you must already grasp the concept dog in order to take yourself to be responding to Fido as you ought. So it would seem like a hopeless prospect to try to explain what it is to grasp the concept dog by appealing to the subject's awareness that she is responding to dogs as she ought. Such awareness, it would seem, presupposes grasp of the concept dog, hence can't explain it.

However it is an essential feature of my proposal that we can make sense of a subject's claim to be behaving as she ought which does not depend on her antecedent grasp of a rule. That is, I want to deny that you have already to have grasped a rule for the use of the word "dog," or for the tokening of the mental representation DOG, in order to take the use of that word or mental representation, to be appropriate in a given situation. The strategy behind my proposal for explaining our grasp of concepts is to reverse the relation that is typically thought to hold between rules on the one hand, and claims to be doing as one ought on the other. Typically it is assumed that a claim to be doing as one ought must rest on a claim to be according with a normative rule. But I am suggesting that we think of the relation between "oughts" and "rules" the other way around. My suggestion in a nutshell is that one counts as grasping a rule in virtue of engaging in a certain course of behaviour with the awareness that one is doing as one ought with respect to one's circumstances, and in particular to what one did previously. And the "ought" here is, as I've put it elsewhere, primitive: one can coherently take oneself to be doing as one ought with respect to one's previous usage without appealing to an antecedently grasped rule which dictates what one ought to be doing. Thus, rather than supposing that we can take ourselves to be doing as we ought in virtue of our

grasp of a rule by which our behaviour is guided, I am suggesting that we count as grasping a rule in virtue of our taking ourselves, in a certain course of behaviour, to be doing as we ought. The notion of primitive normativity, on this suggestion, underwrites that of of a rule: a rule is just something that we grasp in virtue of behaving a certain way with the awareness that we are behaving appropriately in the primitive sense.

I've tried to defend this notion of "primitive normativity" elsewhere: here I'll just try to indicate very briefly a kind of example which might lend it some plausibility. Consider, very intuitively, how we might think of a child acquiring concepts of different geometrical shapes: say cube, cylinder and pyramid. It's very natural to suppose that children acquire these concepts through learning to sort objects of these different kinds. We might imagine a child being presented with blocks of the various shapes, and encouraged, initially by example, to sort them into different boxes. Or we might imagine her learning to sort them by learning the corresponding words: an adult might point to a cube and say the word "cube," and the child might be encouraged to follow her example with other cubes. Now if we think of this as part of what goes into the child's learning a concept, then we cannot say that the child's sorting is being guided by a set of antecedently grasped concepts. If her putting one cube with the other cubes already presupposed her grasp of the concept cube, then her learning to sort the cubes could not contribute to her learning the concept; her putting one cube with the other cubes would at most be a manifestation of her already having grasped the concept. So it cannot be that she sorts as she does as a result of having grasped a concept or rule for sorting: it cannot be that her sorting is guided by some antecedently acquired representation which guides her in putting a new block with the cubes rather than the pyramids.

But at the same time, at least on the face of it, her sorting the blocks seems to have a normative dimension: it seems to involve her own awareness of what she's doing as appropriate. When she puts one cube with the other cubes as opposed to the pyramids, or responds to it with the word "cube" rather than the word "pyramid," she does it with at least an apparent sense that the cube belongs with the other cubes, or that the word "cube" fits the block she is pointing to, given how she used it before, or how the teacher showed her to use it. Her response does not seem to be a mere mechanical reflex: she seems to be manifesting a sense that what she is doing fits her circumstances. And we can imagine her reacting with surprise, or even indignation, to another person who sorts the blocks differently, for example putting a cube

with the pyramids. We can imagine her indicating her sense of the appropriateness of her own response by protesting that the cube "doesn't belong there," but "ought" instead to go in the other box.

I have illustrated the idea of primitive normativity with the case of a child, since that makes it easier to see that one might take one's response to something to be appropriate without first having the corresponding concept. But the suggestion I really want to make is that the notion of responding to one's circumstances as one ought is conceptually prior to the notion of grasping a rule. In other words, the temporal priority I've suggested in the case of a child, between the awareness of responding appropriately and the grasp of a rule, is intended to illustrate a kind of a conceptual or explanatory priority: we can explain, or make sense of, the idea of grasping a rule, in terms of the idea of doing as one ought. So, for example, I want to suggest, we can explain what it is to grasp a rule for discriminating dogs from non-dogs by appealing to the idea of someone's being able to sort dogs from non-dogs with the awareness, in so doing, that she is sorting as she ought. A subject grasps a rule for discriminating dogs in virtue of having the capacity to discriminate dogs, where her discriminative responses involve a sense of their own appropriateness to the objects being discriminated.

It is against the background of this understanding of grasping a rule that I want to identify concepts with rules. You grasp the concept dog in virtue of grasping a rule for discriminating dogs. But grasping a rule for discriminating dogs is, in turn, having a capacity to discriminate dogs with the awareness, in producing each discriminative response, that you are responding appropriately with respect to your previous responses. If the notion of primitive normativity is coherent, then, there is nothing mysterious about grasping a concept: concepts do not, so to speak, inhabit some third realm distinct from our inner mental processes and our overt behaviour. Rather, grasping a concept is just something that we do in virtue of exercising our natural dispositions to sort things in some ways rather than others, but with the sense that what we are doing has a normative dimension.

References

Bach, Kent (2000). Review of Fodor's Concepts. Philosophical Review [check ref.].

Brandom, Robert (2000). Articulating Reasons. Cambridge, Mass.: Harvard University Press.

Byrne, Alex (2005). "Perception and Conceptual Content," in Matthias Steup and Ernest Sosa, eds., Contemporary Debates in Epistemology (Oxford: Blackwell).

Fodor, Jerry (1998). Concepts (Oxford: Oxford University Press).

Fodor, Jerry (2004). "Having Concepts: A Brief Refutation of the Twentieth Century." Mind and Language 19 (1).

Laurence, Stephen and Eric Margolis (2007). "The Ontology of Concepts," *Noûs*, vol. 41.4, pp. 561-93

Margolis, Eric and Stephen Laurence (2005). "Concepts", *The Stanford Encyclopedia of Philosophy (Winter 2005 Edition)*, Edward N. Zalta (ed.)

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

Peacocke, Christopher. A Study of Concepts (1992) (MIT Press, 1992)