Margolis, E. (1998). How to Acquire a Concept. *Mind & Language*, 13:3, pp. 347-369.
DOI: 10.1111/1468-0017.00081

This is a preprint. For the final published version, please visit Mind & Language.

How to Acquire a Concept

Eric Margolis

Abstract: *In this paper, I develop a novel account of concept acquisition for an atomistic theory of concepts. Conceptual atomism is rarely explored in cognitive science because of the feeling that atomistic treatments of concepts are inherently nativistic. My model illustrates, on the contrary, that atomism does not preclude the learning of a concept.*

1. Introduction

In cognitive science, theories of concepts tend to be constrained by an assumption that is so pervasive, it is hardly ever challenged. This is the assumption that what makes a concept the very concept that it is is its relation to certain other concepts.[1] The difference between most theories of concepts consists in the character of the relation they impose. For example, while the classical theory of concepts says a concept C must decompose into a set of concepts that express necessary and sufficient conditions for the application of C, the prototype theory says that C must decompose into a set of concepts that express statistical conditions that govern the application of C. Similarly, the theory-theory—which is gaining attention in psychological circles—says that a concept C must participate in an inferential system of a certain sort and that C is inherently connected to the other concepts that constitute the system.[2] While defenders of each of these theories have emphasized their differences, from the present vantage point, the theories are all strikingly alike. In contrast, one can envision quite a different theory, according to which what makes a concept the very concept that it is is not how it is related to certain other concepts but how it is related to the world. Following current philosophical practice, I'll call theories in this spirit atomistic theories of concepts.[3]

---

[1] The term 'concept' is used in a number of ways in philosophy and cognitive science. Throughout this paper, I'll follow standard psychological usage, according to which concepts are understood as mental representations that are largely individuated by their contents. In contrast, philosophers often insist that concepts are senses—Fregean abstract objects. For some discussion on the relation between these two views of concepts, see Margolis and Laurence, 1998.

[2] For a review of these and related theories, including the exemplar theory, see Smith and Medin, 1981; Komatsu, 1992; Medin, 1989; and Laurence and Margolis, 1998.

[3] Note that, since definitional or statistical analyses can't go on forever, both the classical theory and the prototype theory are committed to the existence of a stock of concepts that aren't themselves subject to analysis. Thus both theories are incomplete, pending an account of what makes these unanalysed concepts the very concepts that they are. Perhaps, at this point, one may want to say that connections to the world supplement the theories—that atomism

Surely one of the more interesting questions in the history of cognitive science is why atomistic theories have received so little attention. This issue is especially pertinent given the difficulties that most theories of concepts are known to have.[4] My own suspicion is that the underlying motivation has always had something to do with the issue of which concepts, if any, are innate. The problem is that, on the face of it, atomistic theories of concepts encourage extreme nativistic positions. This is because models of concept learning generally presuppose that learning a concept is a process in which previously available concepts are assembled in a way that reflects the conditions on the learned concept's identity. For instance, if the classical theory were right, then learning a concept might involve assembling the concepts that express the conditions that are necessary and sufficient for its application. In the same fashion, if the theory-theory were right, then learning a concept might involve learning the theory in which the concept is an essential participant. The trouble with atomistic theories of concepts is that, for any given concept C, there isn't a prescribed set of related concepts whose assembly would constitute having learned C. But if C can't be learned, then presumably it is innate.[5]

Despite the appeal of this line of reasoning, it's a shame that atomistic theories have received so little attention, first, because it's an empirical question which concepts are innate—just because a theory has a strong nativistic commitment doesn't mean we can rule it out from the start—and, second, because other theories of concepts are known to be highly problematic. Who could deny that we are in need of some new theoretical options?

In this paper, I propose to defend atomistic theories of concepts against the claim that they are inconsistent with plausible accounts of concept acquisition. The heart of my defence is a novel model of acquisition that I develop around an atomistic theory, focusing on natural kind concepts. The model I sketch is sensitive to a considerable amount of data that has come to constrain other theories of concepts; yet, it should be said, the data look quite different from the perspective introduced by conceptual atomism. Since any learning model must take into account the nature of what is eventually acquired, I begin, in section 2, by presenting an atomistic theory of concepts—Jerry Fodor's asymmetric-dependence theory. In section 3, I turn to the specific concerns that natural kind concepts raise for the asymmetric- dependence theory. Then, in section 4, I present the model of acquisition. In the first instance, the model is developed around the central case where a concept is acquired by exposure to its instances. In section 5, I add a modification so that it covers the acquisition of concepts that depend upon deference to members of a linguistic community.

2. Concepts, Information-Based Semantics, and Sustaining Mechanisms

Before we can turn to the question of how concepts are acquired within an atomistic framework, we need to address the question of how an atomistic theory of concepts works. It's one thing to

---

is true at the most fundamental level of the conceptual system. Yet defenders of these theories rarely address this issue explicitly, leaving it an open question how exactly the unanalysed concepts are to be treated.

[4] On the classical theory, see Fodor et al., 1980, and Fodor et al., 1975; on the prototype theory, see Osherson and Smith, 1981; Armstrong et al., 1983; Rey, 1983; Margolis, 1994; on the theory-theory, see Margolis, 1995.

[5] An inference that is notoriously endorsed in Fodor, 1981.

say that a concept's identity doesn't depend upon its relations to any other particular concepts; it's quite another to offer a positive account of what the alternatives are. In this section, I explain Jerry Fodor's asymmetric-dependence theory, since it is the most developed atomistic alternative (Fodor, 1990). Because Fodor's theory is a variant on the information-bases semantics approach to content, we should begin with information-based semantics (IBS, for short).

IBS is a schematic semantic theory; it answers the question, What makes a given mental representation have the interpretation or content that it does? For example, what makes it the case that the mental representation bread expresses the property *bread* and not, say, the property *refrigerator*?[6] What unites IBS theories is that their answer to this question depends, in a crucial way, on the information a concept carries, where information is understood in terms of lawful (or reliable) correlations. The number of rings in a cross- section of a tree correlates with the age of the tree, so the rings are said to carry the information about how old the tree is. The expansion of the mercury correlates with the local temperature, so the mercury is said to carry the information about the local temperature. Following these examples, we can say in general that a token or an event carries information about its reliable cause. Put in mental terms what this means is that the concept BREAD expresses the property *bread* because bread is the reliable cause of BREAD-tokenings. In other words, there is a law connecting the property of being bread with the property of being a bread-tokening, and it is in virtue of this law that the concept BREAD expresses the property *bread*.

In the philosophical literature, one difficulty with this sort of account has dominated the discussion—the problem of error. A fundamental constraint on any theory of content is that it must accommodate incorrect applications of a concept. The worry has been that information-based semantics may not be able to meet this constraint. The reason is straightforward. Information- based semantics says that effects carry information about their reliable causes. Yet surely some incorrect applications of a concept are reliable. If at night and at a certain angle and distance a crumpled bag looks like a cat, then you will be bound (at first) to token the concept CAT and not bag. What's more, because under these conditions the bag looks like a cat, the mistake isn't gratuitous. It is a perfectly natural mistake and one that you ought to make not just once, but whenever similar situations arise. After all, bags in these situations look like cats. Here, then, is the problem. The concept CAT has at least two reliable causes. CAT-tokenings are elicited by cats but they are also elicited by crumpled bags. So they carry information about cats and about bags. So why does the concept CAT express the property *cat* and not the disjunctive property *cat-or-bag*? In short, if erroneous applications of concepts can be as reliable as correct applications, an unadulterated information-based semantics has no room for error. There are currently a number of proposals about how to accommodate error. In my view, Fodor's stands out because of the way it locates error within a more fundamental difficulty, which he calls the 'robustness' of meaning.

The problem with erroneous applications of a concept stems from their potential reliability. Yet there are other cases where a concept will be reliably elicited by something other than instances of the property it expresses. Thought is another obvious example. If thinking about fur reliably

[6] To distinguish concepts from properties, I'll adopt a notation where the first are indicated by capitals and the second by italics. Mentioned words are indicated by single quotes.

causes you to think about cats, then (for you) CAT carries information about fur- thoughts. But, of course, CAT doesn't express the property *cat-or-fur-thought*. CAT means *cat*. It just happens that its tokenings are indicative of whether you've been thinking about fur. Information, in other words, is intrinsically tied to etiology; in contrast, the robustness of meaning consists in the distinctness of a symbol's meaning from its cause. All tokens of a symbol mean the same thing however they are caused. Tokens of CAT express the property *cat* whether they are caused by cats, by cat-looking things (in cases of mis- taken identification), or by other thoughts. Hence, if conceptual content is to be constructed out of information, content has to be information plus something else.

Fodor's theory is of special interest since it adds to information in a way that treats error as a special case of the robustness of meaning and not as a unique problem that requires its own, special solution. The theory has two parts:

> (1) A concept stands in lawful relation to the property it expresses.
> (2) Other lawful relations involving the concept are asymmetrically dependent upon the lawful relation between it and the property it expresses.

(1) recapitulates the initial intuition that grounds information-based semantics; concepts express properties they carry information about. However, since they don't always express the properties they carry information about, (2) distinguishes the cases where they do from the cases where they don't. Asymmetric dependence is a relation among laws. Suppose that a concept stands in lawful relations to two properties and hence carries information about both. Let us call these relations L_1 and L_2. Essentially, (2) says that L_2 wouldn't hold but that L_1 holds and not vice versa. L_2 is asymmetrically dependent upon L_1.

The mechanics of the proposal are easiest to see in light of an example. Take the concept CAT again. According to the theory, CAT stands in a lawful relation with the property *cat*, whereby the presence of cats tends to elicit tokenings of the concept CAT. Yet tokenings of the concept CAT may be elicited in other, regular ways, as in the case of the crumpled bag spotted on a dark night. Then, we may suppose, there also exists a lawful relation between the concept CAT and the property *crumpled bag*. However, this lawful relation seems only to hold because the other one holds. If cats didn't cause CAT-tokenings, crumpled bags wouldn't either, but not the other way around. That's why CAT expresses the property *cat*, not the property *crumpled bag*.[7]

Again, this theory differs from most others in philosophy and cognitive science in that mind-world relations do the bulk of the work. On other accounts, conceptual content is largely a matter of how concepts figure in a conceptual system. As a result, other accounts tend to require that people have particular beliefs, or inferential dispositions, that are essential to a concept. In contrast, Fodor's theory says that the beliefs one has are irrelevant to conceptual content as such. To have the concept CAT, one needn't believe anything in particular about cats, so long as the concept stands in the right mind-world relation to the property *cat*. The difference between these

[7] The asymmetric-dependence theory isn't without its problems. I'll discuss several of these in the next section, when we turn to natural kind concepts and the various ways that the theory handles them. For other critical discussions, see the papers in Loewer and Rey, 1991, and Adams and Aizawa, 1994.

two ways of looking at things is made vivid by picturing concepts in terms of a filing system. Think of a concept as a file whose label specifies which concept it is and whose entries count as knowledge structures that, in one way or another, are associated with the concept. The question at stake is what determines how a file is labelled. *Why does one file receive the label 'cat' and another receive the label 'lamppost'?* The usual answer is that the information in the file is essential in determining the file's label. It is because, for example, a file includes specific information about animals and other things that it gets the label 'cat'. The asymmetric-dependence theory, on the other hand, dissociates a file's contents from its label, so the entries of a file are, in a sense, inessential. For Fodor, it is because a file is related to things outside of the filing system (i.e. to properties in the world) that it is labelled as it is (see figure 1).

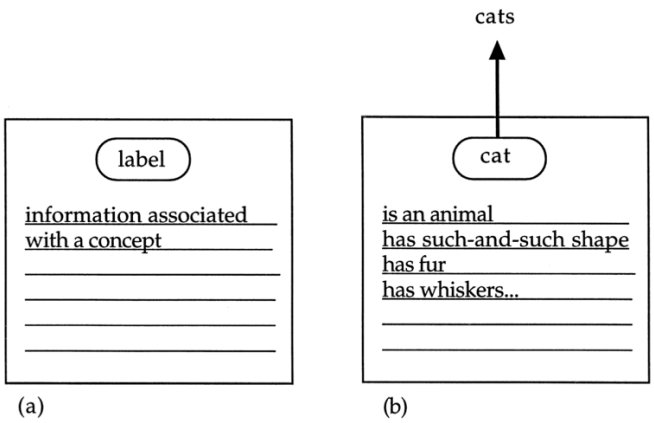

Figure 1. *Concepts may be pictured along the lines of a filing system. (a) represents the general schema. Each file indicates a concept; the entries below a label indicate knowledge structures that are associated with the concept. The question 'What makes a concept express the property it does?' may be recast as the question 'What makes a file have the label it has?' Is it that the file has certain information encoded as an entry, or is it that the file is suitably related to things outside of the filing system? In (b), e.g., the file could receive the label 'cat' because it has information specifically about animals, fur, and so on encoded as entries, or, instead, because it is suitably related to cats.*

But it is one thing to think that a file's contents are inessential to how it is labelled and another to think that they are entirely irrelevant. No one, not even Fodor, thinks they are irrelevant. The reason he doesn't suppose this—indeed, the reason he can't—is that the mind-world relation that he thinks does determine conceptual content must be sustained by a mechanism, and generally the only available mechanism is the inferential apparatus that is associated with a concept. The reason the asymmetric-dependence relation requires a sustaining mechanism is that it isn't at all plausible that the laws in question are basic. Special science laws typically aren't—this is almost a point of definition—and the laws connecting concepts to the properties they express would appear to be on a par with other special science laws, including the laws of psychology.

Now in some cases the sustaining mechanism is going to be noncognitive. This may be true in lower-level perceptual processing, where psychophysical laws are generally expected to have a biological source. But what about the concept TABLE or the concept WALKING or, worse, the concept PROTON? It is hard to see how biology alone could account for the lawful dependence of proton-tokenings on protons, that there is a psychophysical law, as it were, connecting proton with protons. Fodor recognizes this difficulty and suggests putting a twist on a characteristic idea of positivist philosophy of science. The idea, in brief, is to allow that much of what one believes, including the scientific theories one accepts, may be implicated in the mechanism which links, for example, one's PROTON-thoughts with protons. The beliefs a person has endow her with specific inferential dispositions. If someone believes current physical theories and has a sufficient understanding of them, she will have the disposition to infer that a proton is present in certain carefully designed experimental situations where the evidence, in light of what she knows, points to the presence of a proton.

Here's the twist: For Fodor, unlike the positivists, our theoretical beliefs and other knowledge structures aren't constitutive of the concepts whose applications they sustain. So, in particular, one's beliefs about protons aren't constitutive of the concept PROTON. What *is* essential to the concept is that tokens of proton are connected to protons in the way the theory of asymmetric-dependence articulates. As long as tokens of PROTON are suitably connected to protons, it doesn't matter how the connection is sustained (see, for example, Fodor, 1990, note 6, p. 83). You can have your beliefs and I can have mine, and the differences in our beliefs won't in themselves entail that we are subject to conceptual differences. Whether our concepts are different depends upon their connections to the world.

In what follows, I'm going to rely quite heavily on Fodor's version of IBS. This isn't because I think Fodor's theory is entirely correct. Rather, what I find useful about the asymmetric-dependence theory is that it is the most developed and, in my view, the most plausible version of an atomistic theory of concepts. Remember: the question at hand is whether a conceptual atomist can explain the learning of a concept. In this context, Fodor's theory serves as heuristic in that it provides a set of concrete assumptions under which this question can be fruitfully posed.

3. Natural Kind Concepts

To keep the discussion focused, I want to concentrate on just one group of concepts. Partly because they have been prominent in the psychological literature, but also for reasons that will become clear later on, I want to focus on natural kind concepts. Our question, then, isn't how concepts in general are acquired but how natural kind concepts are acquired. And before we can take up the issue of acquisition in earnest, we need to get clear about the way that IBS handles natural kind concepts in an adult's mental life.

It should be clear from the last section that IBS doesn't require of a person who has a natural kind concept that she believes of the things the concept applies to that they are members of a natural kind. (This is for the simple reason that IBS doesn't require of a person that she have any particular beliefs about the things a concept applies to.) Rather, what makes a concept a natural

kind concept is just that it tracks a natural kind in the world; that is, that it is related to a natural kind in the way that IBS requires. Still, this leaves a lot to be said about how it is that our concepts stand in this relation to natural kinds. We need an account of the typical sustaining mechanisms for concepts such as cat and gold and water and proton. These will vary along several dimensions, including the type, amount, and accuracy of the knowledge associated with a concept.

One case that has already come up is where a person actually knows a scientific theory of a kind. Suppose, for example, that a person has had considerable training in physics and chemistry and has assembled a complex set of beliefs about atomic structure, essentially internalizing both the principles of contemporary physical science and the known procedures for manipulating physical particles. Then, because of her hard-earned intellectual resources, she would be in a position to infer the presence of a proton when the available evidence together with what she knows about protons compel the conclusion that a proton is present. In other words, her knowledge puts her in a state of mind where protons cause her to token the concept PROTON, the disposition that is at the heart of the IBS treatment of concepts. Let's call sustaining mechanisms of this kind *theory-based sustaining mechanisms*.

Theory-based sustaining mechanisms clearly have their place in an atom- istic framework. Yet it would be unreasonable to maintain that people always have some theory or other that is rich enough to sustain the right mind-word relation between, as it might be, proton and protons. Another type of sustaining mechanism allows for people to have rather scant knowledge about a kind, so long as they are prepared to exploit other people's more detailed knowledge. I call sustaining mechanisms of this second type *deference-based sustaining mechanisms*. These find their inspiration in Hilary Putnam's proposal that reference depends upon a 'division of linguistic labor' (Putnam, 1975).

Putnam's aim is to rationalize the glaring fact that, for many terms, lots of people are apparently incapable of telling whether they apply in particular cases. He mentions, for example, that he himself is incapable of telling the difference between elms and beeches and hence that he is incapable of telling whether 'elm' or 'beech' applies to a given tree. What are we to make of this and similar examples? Putnam's reaction is to reject, but only partially reject, a verificationist semantics for these terms. He writes (Putnam, 1975, pp. 227–8; emphasis in original):

> everyone to whom gold is important for any reason has to *acquire* the word 'gold'; but he does not have to *acquire the method of recognizing* if something is or is not gold. He can rely on a special sub-class of speakers. The features that are generally thought to be present in connection with a general name—necessary and sufficient conditions for membership in the extension, ways of recognizing if something is in the extension ('criteria'), etc.—are all present in the linguistic community *considered as a collective body*; but that collective body divides the 'labor' of knowing and employing these various parts of the 'meaning' of 'gold'.

In other words, a person may be semantically competent with a range of terms, even if she doesn't know how to tell whether things fall in their extensions, so long as she is suitably related

to people who do. You don't have to be able to tell whether this tree is an elm so long as you can depend upon someone who can—e.g., a botanist. Such people are said to be the experts, and the botanically ignorant are said to use 'elm' deferentially with respect to the experts; that is, to use 'elm' with an intention to be speaking of the things that the experts would identify as elms. In the philosophical literature, terms that are typically used with this sort of intention are said to be deferential terms. These can include natural kind terms, theoretical terms, and perhaps others, such as some social kind terms and artefact terms. By extension, the concepts that these terms encode are said to be deferential concepts.[8]

The notion of a deferential concept is somewhat peculiar. Though it may be true that we have policies about what we take the meaning of a word to be, it is hardly clear that the same thing can be said of the representations in which thinking takes place. Still, there seems to be something right about the appeal to experts in the explanation of concept possession. If we give up Putnam's residual verificationism, a plausible answer is that experts play much the same role as any other sort of evidence. Jerry Fodor has made the comparison especially vivid (Fodor, 1994, pp. 34–5; emphasis in original):

> 'I can't tell elms from beeches, so I defer to the experts.' Compare: 'I can't tell acids from bases, so I defer to the litmus paper'; or 'I can't tell Tuesdays from Wednesdays, so I defer to the calendar.' These three ways of putting the case are, I think, equally loopy, and for much the same reason. As a matter of fact, I *can* tell acids from bases; *I use the litmus test to do so*. And I can tell elms from beeches too. The way I do it is, I consult a botanist. What I do with the litmus, and with the botanist, is this: I construct environments in which their respective states are reliable indicators of the acidity of the sample and the elmicity of the tree; in the one case, I dip the litmus into the fluid, in the other case, I point the expert at the tree. I construct these environments with malice aforethought; with the intention that what color the litmus turns (mutatis mutandis, what the botanist says about the tree) will cause me to have true beliefs about whether the sample is an acid (mutatis mutandis, whether the tree is an elm). In effect, I contrive to replace the problem of determining whether the sample is an acid with the (de facto easier) problem of determining whether the litmus turns red. Likewise, mutatis mutandis, I contrive to replace the problem of determining whether the tree is an elm with the (de facto easier) problem of determining whether the expert calls it one.

In other words, it is simply not true that ordinary people are incapable of telling whether an object is in the extension of a deferential concept. It's just that, in order to tell, they usually need to exploit a special kind of evidence, namely, expert testimony. While Putnam knows next to

[8] Whether deference pertains to just certain classes of terms/concepts or whether it is a more pervasive phenomenon is an interesting question, but one that is irrelevant to the present concern. All that matters is that natural kind terms/concepts are (often) deferential, though I suspect that nearly any term/concept is subject to principles of deference (see Burge, 1979).

nothing about elms and beeches, still he has it within his cognitive resources to discriminate between the two.

The significance of these considerations is that we can imagine a class of sustaining mechanisms for some natural kind concepts which don't require complicated sets of beliefs about the kinds. With deference-based sustaining mechanisms, the agent needs to know some superficial information about a kind (e.g., that its members are called 'elms' in English) and have a disposition to rely upon people who can reliably tell members of the kind from non-members.

So the range of knowledge that goes into a sustaining mechanism for natural kind concepts is quite broad. On the one hand, it could incorporate an internalized body of scientific knowledge and, on the other, it could include as little information as a single yet salient contingent fact about the kind taken together with enough knowledge to locate an expert.

Between these two cases there is another type of sustaining mechanism that is of considerable interest. I call sustaining mechanisms of this third type *syndrome-based sustaining mechanisms*. What I have in mind is a situation where someone, while ignorant of the nature of a kind, nonetheless knows enough contingent information about the kind to reliably discriminate members from non-members without relying upon anyone else's assistance. Take the concept CAT. The case we are imagining isn't one where the person just knows that cats are called 'cat', and it isn't one where the person has a theory of cats, such as a theory about their genetic structure or about the historical facts tying present-day cats to their ancestors. Instead, the proposal is that the person knows about what, for lack of a better term, one might call the cat-syndrome—a collection of salient properties that are readily open to inspection and are reliable indicators that something is a cat. These might include, for instance, the shape of a cat, the typical motions that a cat exercises when it walks or runs, the typical sounds that come out of a cat's mouth, and so on.

Now one of the more significant and problematic features of natural kinds is that their outward, commonly noticed properties aren't constitutive of them. Rather, category membership for natural kinds is determined to a large extent by the hidden, often structural, properties that are responsible for their readily noticed properties. Thus natural kinds are subject to a robust appearance/reality distinction. To take a simple example, a given toy may look tremendously like a skunk, but, even so, it is not a skunk; it's an artefact, a toy. Similarly, a (real) skunk may be altered to look like a cat, but, even so, it remains a skunk and doesn't thereby become a cat. Let's call cases like these *fakes*. A fake is a case where a syndrome that is a reliable indicator of a particular natural kind is instantiated in an item that isn't a member of the kind.

For our purposes, the question is, Given this peculiarity of natural kinds, what sorts of sustaining mechanisms will support a natural kind concept's standing in the mind-world relation that IBS says it does? If we return to the first two types of sustaining mechanisms, fakes don't seem to be a problem. A scientist who has a theory of a kind might be fooled by the appearance of an exemplar, but she would be disposed to correct herself were she to discover that the exemplar lacked the appropriate causal structure of the kind she had mistaken it for. Her theory tells her, for example, that a skunk that is made up to look like a cat isn't really a cat. Similarly, someone who relies upon experts may depend upon people who have such theories. But what about the

case where someone has at her disposal only a syndrome of properties that are indicative of a kind?

The way to handle fakes in cases like these is to grant that, along with the syndrome for a kind, people know something else, some general knowledge that affects their cognitive dispositions, leaving them less governed by the appearance of things. The natural elaboration of this idea is a view that Douglas Medin and Andrew Ortony have aptly called *psychological essentialism* (Medin and Ortony, 1989). According to this view people believe, in general, that category membership for certain kinds of things is determined by the hidden, often structural properties that cause their outward appearances.[9] To be an essentialist is to be disposed to look past the appearance of an exemplar in categorization judgements. The general picture, then, is one where a person's relation to a kind is mediated by two things:

> (1) the person's knowledge of the syndrome for the kind;
> (2) the person's belief that membership within the kind is determined by possession of an essential property (or set of properties) and that this property is a reliable cause of the syndrome.

Together, this information gives the agent the dispositions that asymmetric- dependence requires. Suppose, for example, that the agent were presented with a paradigmatic instance of a cat, where its syndrome is fully apparent. Then the agent would token the concept CAT because her knowledge of the syndrome leads her to infer that she is being presented with a cat. So far, then, we have a mechanism—an inferential mechanism—that mediates the relation between cats and CATS. At the same time, however, the same agent has a disposition to token CAT when presented with a fake cat, e.g., when presented with a skunk which, for one reason or another, looks like a cat. But there is this difference. Were she to find out more information about the latter item—that it lacked the essential property of which the cat-syndrome is a reliable effect—she would cease to apply the concept CAT to it.

Another difficulty that is related to the issue of fakes is owing to a set of examples that have become the focus of attention in philosophical circles — twin cases. Twin cases are instances where two distinct kinds are practically indistinguishable because they fortuitously (yet reliably) exhibit the same apparent characteristics (Putnam, 1975). Here on Earth, H_2O reliably gives rise to the water syndrome, but XYZ, which is located on Twin-Earth, reliably gives rise to the water syndrome too. So a sample of XYZ would cause a normal Earthling to token the concept WATER just as if it were a sample of H_2O. This wouldn't be a problem if it were simply accepted that the concept WATER has both H_2O and XYZ in its extension; then we could say that WATER has a disjunctive essential property. But the intuition that is widely accepted in philosophy is that XYZ falls outside of the extension of water.

I should say that I'm not sure about what to do with twin cases. There are currently a number of proposals in the literature. One is to add an extra clause to the asymmetric-dependence theory, so that reference to the actual world supplements the counterfactuals—since the actual world

[9] The status of psychological essentialism is controversial in psychological circles (Malt, 1994; Braisby et al., 1996), though for reasons that I don't find particularly convincing. For a useful critical discussion of some of the arguments against psychological essentialism, see Abbott, 1997.

contains only H_2O, this may suffice to exclude XYZ (Fodor, 1990). Another is to emphasize the rarity of twin cases (Fodor, 1994). Still another is to argue that twin cases reveal inherent limitations in accounts like the asymmetric-dependence theory, and consequently that they should be abandoned (Adams and Aizawa, 1994). I'm not happy with any of these, but to discuss them all would require a more detailed assessment of asymmetric-dependence—and Twin Earth—than would be warranted here. Instead, what I propose to do is put this difficult issue to the side, except to emphasize one aspect of the way the asymmetric-dependence theory handles twin cases. This is the special case where an agent comes to know that, in fact, there are two essential properties that reliably cause a syndrome.

Take, for instance, the case where a scientist learns about XYZ and comes to distinguish it from H_2O. In this case, although both H_2O and XYZ would cause the scientist to token WATER, the latter is subject to the normal treatment of error. In other words, the XYZ/ WATER connection is asymmetrically-dependent upon the H_2O/ WATER connection. The asymmetry is evidenced by the fact that the scientist, were she to learn that a sample is XYZ, would cease to apply WATER to it, but not the other way around.[10]

4. Acquiring a Natural Kind Concept

Within the IBS framework, acquiring a concept involves establishing a sustaining mechanism that connects the concept with the property it expresses. So, to a large extent, the question of how concepts are acquired amounts to the question of how their sustaining mechanisms are acquired. Now that we've seen what sorts of sustaining mechanisms are involved in the pos- session of natural kind concepts, we are in a position to turn to the issue of how natural kind concepts, in particular, are acquired. In this section, I focus on the central case where, intuitively, it makes sense to say that a concept is learned; that is, where concept acquisition proceeds in the presence of members of the category that the concept picks out. One acquires the concept squirrel and not bee, for example, in the presence of squirrels. This principle isn't universal. Concept acquisition can be facilitated by representations or depictions of the category in the form of pictures, stories, or book learning. What's more, there are any number of cases where acquired concepts refer to objects which, for one reason or another, the subject couldn't interact with. Nonetheless, an account of concept acquisition should be responsive to the fact that the experience leading to the acquisition of a concept is often related to the concept in ways that aren't entirely arbitrary. This is one place where standard, non-atomistic treatments of concept acquisition do well.[11] But

[10] On the other hand, if the scientist doesn't have a disposition to treat H2O and XYZ differently even once she is able to discriminate the two, this is reason to think that, for her, WATER actually applies to H2O and XYZ. In this case, it's the lack of asymmetry that confirms a disjunctive content. Cf. the standard intuition about JADE—an actual case where a single syndrome is caused by two distinct properties (jadeite and nephrite), yet people tend to think that JADE applies to both.

[11] Consider, for instance, the learning model that goes with the prototype theory of concepts. According to the prototype theory, the concept SQUIRREL decomposes into simpler concepts which together express properties that squirrels tend to instantiate. To learn the concept, then, one need only be able to detect the properties and to perform a statistical analysis that keeps track of how properties, in general, tend to cluster. Since the properties cluster in squirrels, it's exposure to squirrels that precedes acquisition of the concept SQUIRREL.

atomistic theories of concepts can also explain the non-arbitrary relation between experience and acquisition, at least when the acquired concept depends upon a syndrome-based sustaining mechanism.

Recall the chief features of a syndrome-based sustaining mechanism. First, the person has to know a collection of salient, relatively accessible properties that are highly indicative of the kind. Second, the person has to have an essentialist disposition; she has to believe that what makes something a member of the category isn't that it exhibits the syndrome but that it possesses the essential property, or set of properties, which is constitutive of the kind and which is a reliable cause of the syndrome. With a little elaboration, the details of this account can be turned into an acquisitional model:

> (1) Young children believe that certain categories are natural kinds and that these categories are subject to a principle of essentialism.
> (2) This principle says that a kind's most accessible properties aren't what determine category membership; rather, it's the possession of an essential property (or set of properties) that reliably causes the syndrome.
> (3) Young children are also predisposed to respond to the types of properties that enter into a kind-syndrome and consequently are highly indicative of a kind.
> (4) There are, in fact, syndromes for some natural kinds.

Assuming all this, we can reconstruct one of the paradigmatic scenarios of concept acquisition, where it's experience of bees, say, that leads to the acquisition of the concept BEE. In the abstract, the way it works is this: The child sees some bees and notices that they have certain properties which suggest that the essentialist principle applies. So the child focuses on the salient accessible properties of the bees and, as a consequence, happens to assemble beliefs about bees that articulate the bee-syndrome. Finally, because of her essentialist disposition, she takes it that something is a member of the kind of which these items are instances so long as it has the same causal structure—the same essential property—which is a reliable cause of these salient and accessible properties. As a result, the child acquires a state of mind—a sustaining mechanism—that links her to bees in the way that the asymmetric-dependence theory requires for her to have the concept BEE. Consider the implications of this mechanism for bees that exhibit the bee-syndrome and for fake bees. Bees that exhibit the syndrome will elicit BEE tokenings simply because they have the right appearance. At the same time, however, fake bees will elicit BEE tokenings too. But, as with the child's adult counterpart, there will be this difference. Were she to find out more information about the fake—that it lacked the essential property that reliably causes the bee-syndrome — she would cease to apply the concept BEE to it.

Of course, from a psychological perspective, this is all very abstract. The model's prospects depends upon, among other things, the plausibility of principles (1)–(4). My own feeling is that they vary in plausibility. Perhaps the most secure is (2). Despite the empiricist tradition in developmental psychology, young children show many signs of an essentialist disposition (for a review, see Gelman and Coley, 1991). First, they appear to be prepared to override gross perceptual similarity in simple induction tasks (Carey, 1985; Gelman and Markman, 1986, 1987). Second, they appear to understand that an object's insides may differ from its outsides and that, for certain kinds of things, the inside of an item is more pertinent to its identity than its

outsides (Gelman and Wellman, 1991). Of course, these data are open to interpretation[12] and more studies need to be done, but, given the present state of the evidence, it's reasonable to conclude that children around the age of 3 or 4 have an essentialist disposition and that this disposition may emerge in children as young as 2. The hard question that this research hasn't settled is, What are the properties that children take to license a special regard for the insides of a novel object? That is, what is it that triggers the essentialist disposition in children for clear cases of natural kinds and not for things like bottles? Also, to the extent that there are syndromes for natural kinds, we need an account of the properties that children are sensitive to such that they do turn out to be reliable indicators of kind membership.

I don't have much to say about the first problem, except to point out that children might depend upon rather coarse heuristics. For example, animate objects are all members of some natural kind or other. So if there were a reliable clue to animacy, this could be used to infer for a range of objects that their insides are especially relevant to their category membership. And, of course, there are some obvious heuristics for identifying animate objects. One of these has to do with their characteristic motion. Unlike chairs or balls, many animate objects don't require an external force to put them into motion. Their movement appears to be spontaneous or internally directed. But what of the syndromes? What are they, and are children really sensitive to the properties that enter into a kind's syndrome? I think a little more can be said on this point.

For some time now cognitive psychologists who have recognized the importance of categorical hierarchies have also recognized a battery of con- verging evidence for distinguishing a level of basic perceptual categories (Brown, 1958; Rosch et al., 1976). The basic level in a taxonomy is the level that, intuitively speaking, is in the middle. For instance, in the hierarchy [*Fido*, *dachshund*, *dog*, *animal*, *physical object*, and *thing*], *dog* marks the basic level; *animal*, *physical object*, and so on are too abstract, and *dachshund, Fido*, and so on are too concrete. Among the chief features of basic perceptual categories is that, in a taxonomic hierarchy, they are the most abstract members whose instances share similar shapes. The notion of shape that's at stake is a matter of controversy. It would be fair to say, however, that all hands agree that a fairly rich notion must be accepted, one that incorporates, for example, the prototypical angle from which the object is viewed. In any event, it is generally recognized that, at the basic level, shape correlates with kind. Thus shape is a prime candidate for a host of kind syndromes.

Moreover, the available evidence suggests that, in certain categorization tasks, young children are guided by their recognition of similarities of shape. The interpretation of these data is highly controversial, but I think a coherent picture is emerging. This takes a bit of explanation, starting with some background on the study of lexical acquisition.

In the past, the typical categorization study in developmental psychology would have children examine an object only to be asked which of two new objects it goes with. The notorious finding for young children has always been that they tend to group objects thematically. For instance, given a base- ball, a volleyball, and a bat, young children might say that the baseball and the bat go together. This choice is considered to be thematic because the ball and bat are related by their typical uses and not by the taxonomic criteria that adults instinctively acknowledge. Under a

[12] For a different perspective and some potentially conflicting data, see Keil, 1989.

taxonomic criterion, the two balls are supposed to go together because they are both the same type of thing: they are both balls.

Recently the conviction that young children's categorizations are dominated by thematic groupings has come under attack. Much of the supporting research for this shift in perspective comes from lexical acquisition studies. The main finding has been that, when young children take themselves to be learning the meaning of a new word, thematic groupings give way to taxonomic ones. Ellen Markman has been at the forefront of this research (for an overview, see Markman, 1989). In a landmark study undertaken with Jean Hutchinson, Markman presented 2- and 3-year-old children with groups of pictures, starting, in each case, with a target picture that was followed by two other pictures. For each triad, one of the follow-up pictures was taxonomically related to the target; the other was thematically related. For example, one triad consisted of a tennis shoe (the target), followed by a high- heeled shoe (the taxonomic choice) and a foot (the thematic choice). The study had two conditions, the 'no word' condition and the 'novel word' condition. The no word condition was just the usual sorting task. The children were asked which of the two pictures goes with the target. The wording used was '. . . See this? [pointing to the target] Find another one that is the same as this . . .'. The novel word condition, on the other hand, involved labelling the target with a noun the children had never heard before. 'See this? [pointing to the target] It's a kind of dax. Can you find another kind of dax?' The results were that children in the no word condition opted for the taxonomic solution on average at a level no better than chance, while children in the novel word condition opted for the taxonomic solution 83% of the time on average (Markman and Hutchinson, 1984).

This and related data have been the source for Markman's advancing the taxonomic assumption, a view about the biases that are inherent to the mechanism responsible for lexical acquisition. The taxonomic assumption says that children assume that a novel word refers to a type of thing rather than a group that is organized thematically. The problem with this proposal—a problem that Markman and her colleagues barely address—is that it can't be taken for granted that children automatically know that two items are of the same type. Indeed, without specific information about how to settle issues of categorical identity, the taxonomic assumption is no better than the economic advice 'buy low, sell high'.

Now, as the evidence for the whole object and the taxonomic assumptions has accrued, so has a body of related evidence for a bias concerning shape. If children act as if they assume that novel words refer to types of things, they also act as if they assume that novel words (in particular, novel count nouns) refer to objects of the same shape (Landau et al., 1988). Thus developmental psychologists have had to consider a second constraint on lexical acquisition, the shape bias.

A common reading of the shape bias is that it competes with the taxonomic assumption and reflects an empiricist conception of children, where children are taken to be perceptually bound. (Shape, on this view, is sup- posed to be a high-level perceptual property of objects.) However, this reading is hardly mandatory. A preferred reading, one that Barbara Landau has stressed, is that the shape bias is in place as a heuristic that supports the taxonomic assumption. That is, while children assume that novel count nouns refer to categories that are organized taxonomically, they depend upon the shapes of objects in making decisions about how to project to new members of a category. The value of this heuristic is that it provides a partial answer to

the difficult question of how children know that two objects are of the same type. As Landau puts it (Landau, 1994, p. 297):

If the object naming system is linked to object shape on the one hand and object kind on the other hand, then young learners might assume that objects of similar shape are also likely to be of similar kind. That is, the links among shape, name, and kind should allow learners to make a critical inference: Objects of similar shape are often also of similar kind.

Landau's suggestion, in other words, isn't that children are subject to the shape bias instead of the taxonomic assumption but rather that the two are integrated. Moreover, Landau takes the shape bias to be just one heuristic among potentially many others. Her proposal is that children have an evolving stack of diagnostics for making category decisions. At the top of the stack is the principle that shape is indicative of category membership, though it isn't decisive and can be overridden by other factors. Indeed, 'In this scheme, most of the burden of development would be placed on learning about other diagnostics for category membership, and thereafter (with increasing attentional and memorial resources), organizing and revising their weighting relative both to shape and to each other' (Landau, 1994, p. 299).

So the stack is subject to change as a child comes to learn more about the sorts of things that count as evidence for category membership, and, as she matures, the expansion of cognitive resources gives her greater facility with the stack and the ability to alter it as her experience dictates. Some of the other diagnostics that Landau seems to think may be on the initial stack include things like texture and function, properties whose salience with respect to shape can be tested in the laboratory. Yet it is clear, on Landau's view, that eventually just about anything can be added as the result of explicit instruction or education. The difference between younger children and older children—or the difference between children and adults, for that matter—is that the older you are, the more beliefs you've accumulated about the properties you take to be diagnostic of category membership.

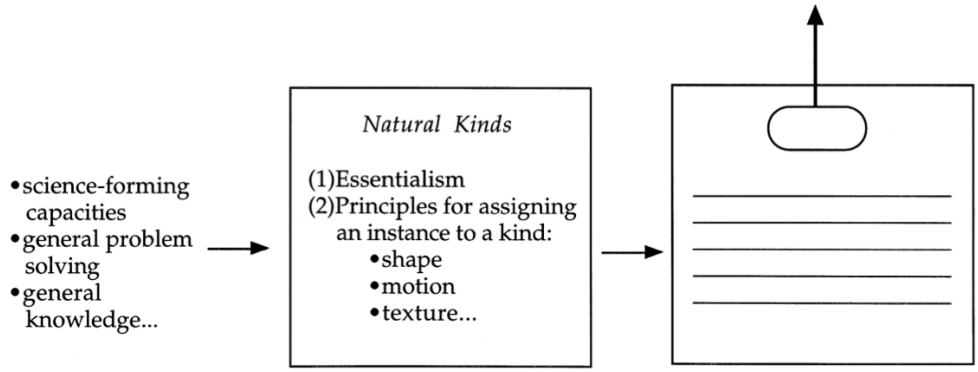

Figure 2. *A model for the acquisition of natural kind concepts. Salient proper- ties of a kind are recorded on the basis of perceptual contact. This information is put in association with a new mental symbol which, in conjunction with the essentialist principle, comes to constitute a sustaining mechanism linking the symbol with the encountered kind.*

Landau's view of lexical acquisition is, I think, loaded with implications for broader issues of cognitive development. This is because her model of lexical acquisition co-opts a relatively general intellectual mechanism. The taxonomic assumption may be a bias of a language learning device, but the heuristics that support it look to be part of a child's understanding of the nature of objects.

If this is right, we may be able to use some of Landau's materials to fill in the gaps in the model of concept acquisition with which we began. The model was supposed to explain the acquisition of natural kind concepts by granting children the ability to accumulate beliefs about the syndromes for natural kinds, where these beliefs would interact with their tacit commitment to essentialism. We've already seen that there is some evidence that children are essentialists. So we've been focusing on the question of whether children are sensitive to the sorts of properties that may be indicative of kind membership. Shape is by far the most convincing candidate. Things like the texture, characteristic local motion, colour, and a host of other properties may be involved as well.

At any rate, we have a first-pass sketch of a model of concept acquisition, and it's one that's responsive to the common situation in which a concept is acquired through experience with an exemplar (see figure 2). According to the model, children have a disposition to acquire natural kind concepts. They understand that certain sorts of things are what they are by virtue of the causal structure that is responsible for their salient and accessible characteristics. They also have a batch of heuristics that are indicative of kind membership, including, importantly, the belief that shape is a good guide to kind membership. And finally, children's heuristics are malleable, changing as they gain more experience and education.

What happens under this model is that a child, in certain situations, takes herself to be confronted by a natural kind. As a result, she records information about the kind that her heuristics tell her is important and puts a record of this information in association with a dummy concept, that is, a previously unused mental symbol. This information together with the essentialist principle comes to constitute a sustaining mechanism which links the symbol with the kind she has encountered. The consequence is that she acquires a disposition under which members of the kind cause tokenings of the symbol. If it's bees she has encountered, the concept she will acquire is BEE; if it's cats that she has encountered, the concept she will acquire is CAT.

5. Concluding Remarks: The Nativism Dispute

At this point, it might prove useful to step back and look at the larger issue of whether the model I've been sketching is implausibly nativistic. The worry for many people has been that if conceptual atomism is correct, then all sorts of concepts that couldn't possibly be innate are innate. Some of the concepts that people keep returning to in the philosophical literature include CARBURETTOR, PROTON, and BROCCOLI—concepts of artefacts, concepts associated with scientific theories, and concepts of ordinary natural kinds. Again, I should say that I'm not

sympathetic with the claim that these concepts must be learned, since the innate structure of the mind is clearly an empirical question; and elsewhere I've argued against the chief reason that philosophers have offered for supposing that broccoli and its ilk couldn't be innate; namely, that evolutionary theory wouldn't allow it (Margolis, 1998).

Still, to the extent that there is something to the empiricist intuition, the present model of concept acquisition may suffice to temper some of the resistance to conceptual atomism. This is because the model doesn't entail that natural kind concepts themselves are innate. Rather, individual natural kind concepts are learned by exploiting a mechanism that is responsible for the acquisition of a whole class of concepts.

To see how this works, consider the question of why the child learns the concept BROCCOLI, given the model as it has been developed so far. It's not because the concept BROCCOLI is innate, just waiting to be elicited. Rather, it's because of two things. First, the child has a more general intellectual endowment—perhaps an innate intellectual endowment—that suits the purpose of acquiring natural kind concepts. Second, the child has experiences with broccoli. If she had experiences with a different kind, she would have acquired a different syndrome and hence a different concept. Thus the model has a distinctive Kantian flavour. It explains the acquisition of a range of concepts against the background of a disposition to view the world in terms of a distinctive human category.

Notice that this sort of explanation is quite different from the one that is usually associated with atomistic theories of concepts. Atomistic theories are generally thought to entail that concepts can only be acquired by a triggering process. This means, among other things, that unstructured concepts are innately specified, only to be made available to cognition when an innately specified 'trigger' is encountered. In contrast, the model I've been encouraging is hardly a triggering-theory. There may be constraints on which concepts can be acquired, but they are rather general constraints; they guide the acquisition of a certain type of concept. Moreover, specific concepts are acquired in a way that is responsive to the experience that the agent undergoes. It's not just that experience is a precursor to concept acquisition. Experience is crucial to the development of the syndromes that come to support concept possession. As a result, the model of acquisition has the spirit of a learning model. It explains how a concept, of a certain type, might be learned in the course of experience, even if it lacks internal structure.[13]

One of the virtues of this account is that it doesn't require that children know too much before acquiring a natural kind concept. Most importantly, they don't have to know the nature of a kind. What they have to know is some contingent information about a kind, information that is available given perceptual contact with one or more of its exemplars. But, while perceptual contact is often a precursor to concept acquisition, it isn't a precondition. Certainly, people have the ability to acquire natural kind concepts without perceptual contact, and there is no reason to think that this ability is limited to an achievement of adult life. Here it pays to bear in mind one

[13] Natural kind concepts constitute one type to be guided by a special acquisitional mechanism. Perhaps there are other types, and corresponding mechanisms, as well—a question I'll leave for another time.

of the sustaining mechanism types that, with adults, allows for the possession of a natural kind concept in the face of gross ignorance—deference-based sustaining mechanisms.

Adults, we've seen, can exploit expert testimony, and consequently have the ability to discriminate members of a kind of which they know almost nothing. All they do have to know is some small amount of contingent information about the kind, and some rather general information about how to locate a suitable expert. But if this is how it is for adults, why not for children? We can just add to Landau's stack of heuristics an entry that says to defer to expert testimony. Of course, children may not know who the experts are. That is something to be learned in the course of development. However, even if children start by assuming that their caretakers are the experts, this may do a lot to extend their conceptual repertory. Though the caretakers may not themselves be experts, they might be able to locate the real experts and consequently act as a medium of reliable discrimination. After all, the caretakers will be adults and will have acquired the experience and knowledge that's common within society.

What's more, children probably do have the disposition to rely upon adult testimony. One of the things that causes them to override perceptual similarity in an induction task is that the perceptually dissimilar items are referred to by the same name. By hypothesis, the reason children conclude that the dissimilar items have the same hidden properties is that they are following the taxonomic assumption and applying it backwards: knowing that two items have the same name, children infer that they are of the same kind. But children aren't forced to accept this conclusion. They could infer, instead, that the experimenter or the puppet or whoever is talking is using the names incorrectly. Deference, then, can be used to supplement the core model by providing an entirely general disposition that helps to put in place specific sustaining mechanisms. Moreover, since deference may not be restricted to natural kind terms/concepts (see note 8), it may offer a way of explaining the acquisition of nearly any lexicalized concept where the child lacks sufficient experience of its instances.

In short, conceptual atomists have a number of resources to explain how concepts are learned. In each case the key to acquisition is a process that constructs a sustaining mechanism that effects the mind-world relation that is constitutive of conceptual identity. This orientation shows that one of the main arguments for rejecting atomistic theories of concepts is simply unfounded. Unstructured concepts needn't be innate.

References

Abbott, B. 1997: A Note on the Nature of 'Water'. *Mind*, 106, 311–19.

Adams, F. and Aizawa, K. 1994: Fodorian Semantics. In S. Stich and T. Warfield (eds), *Mental Representation: A Reader*. Cambridge, MA: Blackwell, 223–42.

Armstrong, S., Gleitman, L. and Gleitman, H. 1983: What Some Concepts Might Not Be. *Cognition*, 13, 263–308.

Braisby, N., Franks, B. and Hampton, J. 1996: Essentialism, Word Use, and Concepts. *Cognition*, 59, 247–74.

Brown, R. 1958: How Shall a Thing Be Called? *Psychological Review*, 65, 14–21.

Burge, T. 1979: Individualism and the Mental. In P. French, T. Uehling, Jr and H. Wettstein (eds), *Midwest Studies in Philosophy, IV: Studies in Metaphysics*. Minneapolis: University of Minnesota Press, 73–121.

Carey, S. 1985: *Conceptual Change in Childhood*. Cambridge, MA: MIT Press. Fodor, J. A. 1981: The Present Status of the Innateness Controversy. In *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press, 257–316.

Fodor, J. A. 1990: A Theory of Content. In *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press, 51–87.

Fodor, J. A. 1994: *The Elm and the Expert: Mentalese and Its Semantics*. Cambridge, MA: MIT Press.

Fodor, J. A., Garrett, M., Walker, E. and Parkes, C. 1980: Against Definitions. *Cognition*, 8, 263–367.

Fodor, J. D., Fodor, J. A. and Garrett, M. 1975: The Psychological Unreality of Semantic Representations. *Linguistic Inquiry*, 6, 515–32.

Gelman, S. and Coley, J. 1991: Language and Categorization: The Acquisition of Natural Kind Terms. In S. Gelman and J. Byrnes (eds), *Perspectives on Langu- age and Thought: Interrelations in Development*. Cambridge University Press, 146–96.

Gelman, S. and Markman, E. 1986: Categories and Induction in Young Children. *Cognition*, 23, 183–209.

Gelman, S. and Markman, E. 1987: Young Children's Inductions from Natural Kinds: The Role of Categories and Appearances. *Child Development*, 58, 1532–41.

Gelman, S. and Wellman, H. 1991: Insides and Essences: Early Understandings of the Non-Obvious. *Cognition*, 38, 213–44.

Keil, F. 1989: *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT Press.

Komatsu, L. 1992: Recent Views of Conceptual Structure. *Psychological Bulletin*, 112, 500–26.

Landau, B. 1994: Object Shape, Object Name, and Object Kind: Representation and Development. In D. Medin (ed.), *The Psychology of Learning and Motiv- ation*. New York: Academic Press, 253–304.

Landau, B., Smith, L. and Jones, S. 1988: The Importance of Shape in Early Lexical Learning. *Cognitive Development*, 3, 299–321.

Loewer, B. and Rey, G. (eds) 1991: *Meaning in Mind: Fodor and His Critics*. Cam- bridge, MA: Blackwell.

Laurence, S. and Margolis, E. 1998: Concepts and Cognitive Science. In E. Marg- olis and S. Laurence (eds), *Concepts: Core Readings*. Cambridge, MA: MIT Press.

Malt, B. 1994: Water Is Not H_2O. *Cognitive Psychology*, 27, 41–70.

Margolis, E. 1994: A Reassessment of the Shift from the Classical Theory of Concepts to Prototype Theory. *Cognition*, 51, 73–89.

Margolis, E. 1995: The Significance of the Theory Analogy in the Psychological Study of Concepts. *Mind and Language*, 10, 45–71.

Margolis, E. 1998: Why Evolution Can't Settle the Nativism Controversy. Under review.

Markman, E. 1989: *Categorization and Naming in Children: Problems of Induction*. Cambridge, MA: MIT Press.

Markman, E. and Hutchinson, J. 1984: Children's Sensitivity to Constraints on
Word Meaning: Taxonomic vs. Thematic Relations. *Cognitive Psychology*, 16, 1–27.

Medin, D. 1989: Concepts and Conceptual Structure. *American Psychologist*, 44, 1469–81.

Medin, D. and Ortony, A. 1989: Psychological Essentialism. In S. Vosniadou and
A. Ortony (eds), *Similarity and Analogical Reasoning*. Cambridge University Press, 179–95.

Osherson, D. and Smith, E. 1981: On the Adequacy of Prototype Theory as a Theory of Concepts. *Cognition*, 9, 35–58.

Putnam, H. 1975: The Meaning of Meaning. In K. Gunderson (ed.), *Language, Mind and Knowledge*. Minneapolis: University of Minnesota Press. Reprinted in H. Putnam (1975), *Mind, Language and Reality: Philosophical Paper*, vol. 2. Cambridge University Press, 215–71.

Rey, G. 1983: Concepts and Stereotypes. *Cognition*, 15, 237–62.

Rosch, E., Mervis, C., Gray, W., Johnson, D. and Boyes-Braem, P. 1976: Basic Objects in Natural Categories. *Cognitive Psychology*, 8, 382–439.

Smith, E. and Medin, D. 1981: *Categories and Concepts*. Cambridge, MA: Harvard University Press.