WHERE CONCEPTS COME FROM: A THEORY OF CONCEPT ACQUISITION

William Dylan Sabo

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy.

Chapel Hill
2008

Approved by:
Jesse Prinz
William Lycan
Ram Neta
Dorit Bar-On
John Roberts
ABSTRACT

WILLIAM DYLAN SABO: Where Concepts Come From: A Theory of Concept Acquisition
(under the direction of Dr. Jesse Prinz)

How do people acquire new concepts? Most theorists (including Quine, Chomsky, Fodor, and many others) assume that childhood learning is a kind of theory-building. This picture implies that children acquire new concepts by deploying concepts they already possess, and that in turn implies what I call the Conceptual Mediation Thesis (CMT): that, in order to acquire any new concepts, a cognizer must first already have some concepts.

I argue that CMT is false. While CMT implies that at least some concepts are innate, it is widely accepted because it is thought to provide the only way to explain how concepts are acquired. However, I argue that the apparent explanatory virtues of CMT are in fact illusory. I then show how we can satisfy the explanatory goals that CMT was supposed to satisfy without postulating any innate concepts – indeed without any innate mental representations at all.

I distinguish between indicating states and representing states of cognizers. Indicating states differ from representing states in being stimulus-bound: only those tokens directly caused by what they indicate count as correct. I argue that perception produces states that indicate features of the environment. These indicating states serve as input to mechanisms that record these states. These recording devices, in turn, respond to
the input of systematically similar indicating states by creating states that represent what those indicating states merely indicate. I describe some processes whereby these recording devices can create representational states without any representational input. I argue that this explanation requires no appeal to mental representations that the agent already possessed. Finally, I show that this approach to concept acquisition has the resources to explain a variety of psychological phenomena that traditional views struggle to accommodate.
For Mom
1950-2004
ACKNOWLEDGEMENTS

I would like to thank Jesse Prinz both for his unfailing encouragement and for his discerning and insightful comments and criticism on version after version of this dissertation. I would also like to thank Dorit Bar-On, Bill Lycan, Ram Neta, and John Roberts for comments, questions, suggestions and support for all of the material here, as well as many previous versions of that material. Finally, I would like to thank Bryce Huebner, Eric Mandelbaum, Clair Morrissey, Mark Phelan, and David Ripley for many helpful discussions of the issues involved. Without them this would not be the work that it is.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................................................ ix

INTRODUCTION ........................................................................................................................................... 1

CHAPTER

1 CONCEPT ACQUISITION AND CONCEPT NATIVISM .............. 6

1.1 Introduction ...................................................................................................................................... 6

1.2 What Concepts Are: Preliminary Considerations ...................... 8

1.3 Concept Learning And The Argument For Radical Concept
    Nativism ........................................................................................................................................ 18

1.4 The Conceptual Mediation Thesis And The Representational
    Necessitation Thesis .................................................................................................................. 29

1.5 Conclusion ...................................................................................................................................... 34

2 NATIVISM, CONCEPT ACQUISITION, AND PSYCHOLOGY ........ 36

2.1 Introduction ...................................................................................................................................... 36

2.2 Methodological Nativism ........................................................................................................... 38

2.3 Nativism, Computationalism, And Psychological Explanation ........ 45

2.4 Conclusion ...................................................................................................................................... 56

3 NATIVIST VIEWS OF CONCEPTUAL DEVELOPMENT .......... 58

3.1 Introduction ...................................................................................................................................... 58

3.2 The Triggering Conception Of Concept Acquisition ............... 59

3.3 Problems With Triggering ......................................................................................................... 62
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Fodor’s Mind-Dependence Proposal</td>
<td>81</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusion</td>
<td>85</td>
</tr>
<tr>
<td>4</td>
<td>RECENT APPROACHES TO EXPLAINING CONCEPT ACQUISITION</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>87</td>
</tr>
<tr>
<td>4.2</td>
<td>Conceptual-Role Semantics And Concept Acquisition</td>
<td>88</td>
</tr>
<tr>
<td>4.3</td>
<td>Sustaining Mechanisms And Concept Acquisition</td>
<td>125</td>
</tr>
<tr>
<td>4.4</td>
<td>Conclusion</td>
<td>135</td>
</tr>
<tr>
<td>5</td>
<td>CONCEPT ACQUISITION WITHOUT REPRESENTATIONAL MEDIATION</td>
<td>117</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>117</td>
</tr>
<tr>
<td>5.2</td>
<td>Representation And Indication</td>
<td>118</td>
</tr>
<tr>
<td>5.3</td>
<td>Perception And Representation</td>
<td>125</td>
</tr>
<tr>
<td>5.4</td>
<td>Concept Acquisition Without Representational Mediation</td>
<td>135</td>
</tr>
<tr>
<td>5.5</td>
<td>Explanatory Goals Of Theories Of Concept Acquisition</td>
<td>149</td>
</tr>
<tr>
<td>5.6</td>
<td>Empirical Support</td>
<td>152</td>
</tr>
<tr>
<td>5.7</td>
<td>Conclusion</td>
<td>162</td>
</tr>
<tr>
<td>6</td>
<td>CONCEPTUAL REPRESENTATION AND CONCEPTUAL MEDIATION</td>
<td>164</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>164</td>
</tr>
<tr>
<td>6.2</td>
<td>Conceptual And Nonconceptual Mental Representations</td>
<td>165</td>
</tr>
<tr>
<td>6.3</td>
<td>Concept Acquisition Via Conceptual Mediation</td>
<td>176</td>
</tr>
<tr>
<td>6.4</td>
<td>Conclusion</td>
<td>185</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>187</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1 .......................................................................................................................... 134
2 .......................................................................................................................... 136
3 .......................................................................................................................... 138
4 .......................................................................................................................... 139
5 .......................................................................................................................... 141
6 .......................................................................................................................... 144
INTRODUCTION

Where do concepts come from? John Locke (1690/1975) answered “in one word, from experience”. But the contemporary consensus in cognitive science has been less friendly to the Lockean position. This contemporary consensus is decidedly nativist in its orientation; according to it, a large amount of conceptual structure is required in order for us to acquire any of our concepts from experience at all. At a minimum, this consensus is reflected in a commitment accepted by virtually all extant theories of concept acquisition to what I call here the Conceptual Mediation Thesis (CMT): that the acquisition of any concept must be mediated by concepts one already possesses. This requires that we have concepts that we do not acquire –concepts we have innately. So, according to current consensus, we need innate concepts before we can acquire any others. Furthermore, the contemporary consensus holds, the number of innate concepts we need is quite large. So, contra Locke, a great many of our concepts are not acquired from experience.

I think that Locke was much nearer to being right about the origin of our ideas than it has been recently popular to suppose. I think it is both coherent and plausible to think that how we acquire our initial conceptual repertoire can be explained without supposing that we have any innate concepts at all. On my view, as on Locke’s, experience is the origin of our ideas, and the possibility of such experience does not require that we have any ideas innately.
Despite my sympathy for the spirit of Locke’s position, however, I do not endorse every aspect of his view. To anticipate, my view differs from Locke’s in at least the following ways:

First, Locke did not distinguish between impressions and ideas. Because of this, Locke was unable to explain the difference between how things appear to us to be and how we judge them to be. Similarly, he was unable to explain how someone could have an experience, in any sense at all, without having ideas, since, on his view, to have an experience simply is to have ideas.

What is needed, to satisfy these explanatory ambitions, is to enforce a distinction between mental representations and more primitive kinds of contentful mental states. Hume provided such a resource in introducing a distinction between impressions and ideas, which allowed him to offer an explanation of how the latter are acquired on the basis of the former.

In my account, this distinction between impressions and ideas becomes the distinction between indicating and representing states. The capacity to have certain kinds of impressions is, for Hume, something we are born with, but that does not mean that we are born with any ideas. Similarly, I argue, we are born with the capacity to have certain kinds of indicating states, in virtue of the way our perceptual systems are built, but that does not mean that we are born with the ability to mentally represent anything. Mental representation, and especially conceptual representation, is a later sophistication that develops from indication but is not identical to it.

Second, Locke, like his contemporaries, relies on introspection to explain the content of ideas and the relations between them. But this is an aspect of his view that
should be rejected. We do not now suppose that that the processes that underwrite psychological phenomena are manifest to consciousness. In explaining how concepts are acquired I have appropriated the term “Similarity” as a general term for relations between various features of our experiences. But I do not think that what relations those are, nor how our cognitive subsystems use those relations, is manifest to introspection. In the same vein, what are the basic data for perception is a question the answer to which cannot be established introspectively. For just this reason, I do not assert that the initial indicating abilities we have must be some particular kinds. The claim that we have them at all is itself meant as an empirical claim (for which, however, I do think there is very good evidence).

Third, Locke, like his fellow empiricists, viewed perception as merely a sequence of sensations. But, as Rationalist critics of Locke argued, this is an inadequate view of perception. Perception does not consist merely in having a sequence of sensations; instead, perception is a matter of having sensations that are structured according to rules. But this reliance on rules, these critics allege, requires that mental representations be involved in perception.

On my view, perception is structured, not by innate representations of rules, but by the ways in which our perceptual and cognitive systems are built to track and acquire information about the world. This is in a way a Kantian view, in that it appeals to various categories to explain how we perceive a world of three-dimensional objects causally interacting in space and time. But, on my view, these categories are not represented, nor do they need to be in order for them to be effective in structuring our perceptions. They
structure our perception of the world by imposing constraints on the operations of our perceptual systems, not by being represented in the operations of those systems.

Fourth, Locke was an empiricist: he held that all mental representation is ultimately perceptual in nature. On this issue I take no stand here. While my view is that mental representation has its origin in perception, I remain neutral on the question of whether that representation has a distinctively perceptual content or character.

Fifth, and most importantly, I think Locke’s account of the mechanisms by which we acquire our ideas is mistaken. Locke – and he has been followed in this by nearly everyone working on this topic – distinguished between simple and complex ideas. On Locke’s view, some of our ideas (Locke thought these were the ideas of sensory qualities) are primitive, and the rest are constructed out of those primitive ideas. But Locke offered no explanation for how the primitive ideas were acquired through experience. Because of this, it has been standard to assume that the only way in which concepts could be acquired is by constructing them from other concepts. This view is one manifestation of commitment to CMT.

The influence of this aspect of Locke’s view has been so pervasive as to render alternatives to it almost inconceivable. It might seem, and has seemed to many, that commitment to CMT is just a condition on what it is to be a theory of concept acquisition. Accordingly, my purpose in the greater part of this work is to loosen the grip that CMT has. I argue that, far from being an uncontroversial condition on what it is for something to count as a theory of concept acquisition at all, it embodies an extremely contentious and severely blinkered view of the space of possibilities for explaining where concepts come from.
In chapter 1 I explain the considerations that have motivated commitment to CMT, and trace out how these considerations have led many to draw nativist conclusions from them. In chapter 2 I explore how these considerations are supposed to constitute arguments for nativism, and argue that it is only a prior commitment to CMT that makes it seem as though CMT is a constraint on an adequate theory of concept acquisition.

Having addressed such dialectical issues, I turn in chapter 3 to an evaluation of nativist accounts of conceptual development, specifically to the conception of conceptual development as a matter of being triggered by experience. I argue that such views have problems satisfying the explanatory goals that a theory of conceptual development should meet. In chapter 4 I critically discuss recent theories of concept acquisition that embrace a commitment to CMT, and argue that they have problems satisfying such explanatory goals as well. These discussions pave the way for an alternative theory of concept acquisition that is not committed to CMT.

This alternative theory of concept acquisition is the subject of chapter 5. On this view, we acquire our initial concepts through experience, in ways that do not presuppose that we already have other concepts. Chapter 6 addresses some remaining issues, including the distinction between nonconceptual and conceptual mental representations, and how to understand the phenomenon of using concepts to acquire other concepts.

I have here indicated several inadequacies in Locke’s view of how we acquire our concepts, and I have indicated my own preferred responses to these problems. But I maintain that my approach here remains very much in the spirit of Locke’s guiding conviction: that the contents of the mind have their source in experience.
CHAPTER 1  CONCEPT ACQUISITION AND CONCEPT NATIVISM

1.1 Introduction

Nativist hypotheses have informed research in cognitive science since its beginning. Noam Chomsky (1959, 1965, 1968) both rejected behaviorism and presaged the “cognitivist revolution”, in the process legitimizing the appeal to mental states and processes in cognitive psychology. Chomsky argued for reconceptualizing psychological research as a matter of tracing the role of experience in operating on the innate endowment of the child, resulting in a mature competence with some cognitive capacity.

Chomsky’s model of developmental explanation has three parts. First is a characterization of the mature competence of a cognizer; this is the end state in which development results. The second is a characterization of the role of experience in developing that mature competence. The third is a characterization of the initial state of the cognizer, which combines with the experiential input to produce the mature capacity. The level and kind of sophistication of the initial state can be determined by first determining the character of the mature competence and then subtracting the contribution made by experience. What cannot be accounted for in the mature competence by the experiential contribution must be a part of the initial state. And since, by hypothesis, the nature of the initial state cannot itself be determined by experience, the initial state must be considered part of the cognizer’s innate endowment.
While there are debates over how to characterize the mature competence of cognizers in various different areas of cognition, as well as over the role that experience plays in cognitive development, this general schema is widely accepted. And acceptance of this general schema fuels flourishing research programs in various aspects of developmental psychology.

This acceptance raises a further question about how to characterize the initial state. While it is clear, on this picture, that some things must be innate, it is not as clear what kinds of things must be. What sorts of states and processes must be innate? In particular, does adherence to this kind of explanatory schema require innate concepts? Many have thought that it does. For it has seemed to many that explaining our conceptual competence requires appeal to a basic set of primitive concepts, with the acquisition of other concepts explained in terms of operations on that primitive base. The question then is about which concepts must be regarded as being in the primitive conceptual base, and how many might be needed in order to explain our full competence. Much research in developmental cognitive psychology, in consequence, has focused on the ways in which we use our primitive conceptual base to learn new concepts.

But this approach to explaining concept acquisition has a surprising consequence. For there are arguments – presented most explicitly by Jerry Fodor (1975, 1980, 1981, 1998) – that purport to show that concepts cannot be acquired in this way. When combined with the claim that this is how concepts must be learned, the effect of these arguments is to motivate a radical nativism about concepts, according to which concepts simply cannot be acquired at all.
In this chapter I will explain the challenge posed by arguments for radical concept nativism. I will trace the motivation for these arguments to acceptance of what I call the Conceptual Mediation Thesis.\(^1\) I will explain this thesis and the role it plays in the debates over concept acquisition. I think acceptance of CMT is at the root of the motivation for radical concept nativism, and that rejecting CMT is a necessary step towards an adequate theory of concept acquisition. So my negative project in this work is to explain why CMT should be rejected. My positive project is to provide a guide to how it can be.

1.2 What Concepts Are: Preliminary Considerations

In order to have a theory of the processes by which concepts are acquired, we need to understand the end state in which such processes are supposed to eventuate. We need a view about what concepts are. But there is little consensus over the answer to this question, and the problem is exacerbated by the fact that different theories often seem to be characterizing different theoretical categories. This raises the worry that “concept” may ultimately not be a term of theoretical interest at all. So it may seem as though, before we can address the question of how concepts are acquired, we need to develop and defend a particular theory of the nature of concepts.

I think we can circumvent these worries by developing a functional profile of concepts. There are some widely agreed-upon explanatory goals that concepts are

---

supposed to meet. We can use these to give a characterization of the explanatory profile of concepts. So while we cannot give an uncontroversial characterization of the nature of concepts, we can provide a characterization of what concepts are supposed to do. Once we have such a “task analysis” for concepts, we will be able to raise the question of how cognizers can acquire abilities to perform these tasks. Thus a functional profile of concepts will exercise a constraint on theories of concept acquisition, without requiring a previously articulated theory of the nature of concepts. In this section I articulate some widely accepted conditions on the nature of concepts, in order to give a preliminary characterization of the end state in which concept acquisition eventuates.

1.2.1 Semantic Features of Concepts

One central role that concepts are invoked to fill is in an explanation of how it is that a cognitive system has the capacity to think. According to the widely-accepted Representational Theory of Mind (Fodor 1975, 1987, 1998; Sterelny 1990), a cognizer’s having a thought is a matter of that her standing in an appropriate relation to a contentful item; that is, an item with conditions of semantic evaluation. For a person to think that-p is for her to stand in some appropriate relation to a mental representation with the content that-p. The Fregean idea of thinking as a matter of grasping a proposition is thus explained by such theories in terms of bearing an appropriate relation to a mental representation that, in turn, stands in a representational relationship to its intentional object.

According to RTM, the mental representations that are the objects of propositional attitudes are structured objects. This is because of certain features thinking
seems to have. First, thinking seems to be *productive*: human cognizers are apparently able to entertain indefinitely many different distinct thoughts. For example, a cognizer with the capacity to think that-\(p\) and the capacity to think that-\(q\), and the capacity to join different thoughts together in logical relations, thereby has the capacity to think that-\(p\)-and-\(q\), that-if-\(p\)-then-\(q\), that-if-\(p\)-then-(\(p\) and-\(q\)), and so on. Second, thinking seems to be *systematic*: the ability to think one kind of thought seems to bring with it the ability to think various other related thoughts. For example, a cognizer with the capacity to think the thought \(\text{ALICE ADORES QUINCY}\) would also seem thereby to have the capacity to think \(\text{QUINCY ADORES ALICE}\), and a cognizer that additionally had the capacity to think \(\text{JAVIER DESPISES PAULINE}\) would seem thereby to have the capacity to think \(\text{ALICE DESPISES QUINCY}\), or \(\text{PAULINE ADORES ALICE}\).²

These features of thinking suggest that mental representations are *compositional*: that the content of a given mental representation that functions as the object of a propositional attitude is a function of its constituent parts. In some cases, these parts will include other mental representations that can function as the objects of propositional attitudes. For example, the object of the thought \(\text{DOGS BARK AND CATS MEOW}\) has as constituents \(\text{DOGS BARK}\) and \(\text{CATS MEOW}\), and each of these in turn is able to function as the object of a propositional attitude. One can simply think \(\text{CATS MEOW}\), for example. But explaining productivity and systematicity also seems to require constituents of mental representations that *cannot* themselves be objects of propositional attitudes. Explaining the relationship between the thought \(\text{ALICE ADORES QUINCY}\) and the thought \(\text{QUINCY ADORES ALICE}\) seems to require us to posit mental representations of Alice, of

² I will use small capital letters for names of mental representations, and italics for the contents of those representations (as well as for emphasis).
Quincy, and of adoring. But none of these can by themselves be the objects of propositional attitudes; one can’t have the thought ALICE, for example. So explaining thinking as compositional seems to require the postulation of mental representations that are the analogue of subsentential expressions in thought. Concepts are the mental representations that serve this function.

So explaining thinking by appeal to mental representation requires the postulation of concepts as the constituent parts of thoughts. Concepts are therefore items with conditions of semantic evaluation themselves. They contribute their semantic properties to the complex mental representations in which they occur, and thus provide propositional objects with their truth-conditions as a function of the satisfaction-conditions of the concepts that compose that propositional object.

It is a further question how the semantic properties of concepts are fixed. There are two broad approaches to this question. According to the first, a concept’s conditions of semantic evaluation are determined by some kind of relation that the concept stands in to the feature of the world that it represents. Within this category, there are various candidates for the content-constitutive mind-world relation. On some accounts, the relation is causal; on others, the relation is informational; on still others, the relation is teleological. These theories concern the relation that must obtain between a mental representation and the aspect of the world that it represents in order for the former to represent the latter.

According to the second approach, a concept’s semantic properties are determined by the relations that it bears to other concepts. As before, there are variations within this

---

3 For causal theories of content, see Kripke (1972), Putnam (1975); for informational theories, see Dretske (1981), Fodor (1990); for teleological theories, see Dretske (1988), Millikan (1984, 1989).
category, but the common thread is that the content of one concept is not independent of the content of other concepts. Usually these relations are inferential relations; concepts, on such views, are essentially related to other concepts (perhaps to perceptual inputs and behavioral outputs as well).⁴

Regardless of which approach one favors to the contents of concepts, one can agree that the fact that a concept has the conditions of semantic evaluations it does is central to its being the concept that it is, and that concepts are essentially items with representational contents. One central explanatory role for concepts, then, is to explain the representational features of thought.

1.2.2 Concepts and Psychological Explanation

In addition to explaining the contents of thought, concepts play a central role in explaining behavior. The reason for this parallels the reason given above for postulating concepts to explain the semantics of the propositional attitudes. Just as we need mental representations to explain the intentionality of propositional attitudes, and we need concepts to explain such representations, so we need propositional attitudes to explain behavior, and we need concepts to explain the propositional attitudes. Just as concepts have semantic properties in virtue of which the objects of the propositional attitudes in which they occur have their semantic properties, so concepts have causal properties in virtue of which the objects of the propositional attitudes in which they occur exercise control over behavior. We act in the ways that we do because of the beliefs and desires that we have, and what beliefs and desires we have is constrained by what concepts we have.

⁴ These two approaches need not be exclusive. One could hold that both mind-world relations and concept-concept relations jointly determine a concept’s semantic properties. I will consider hybrid views in more detail in chapter 4.
have. If you have no CRÈME BRULEE concept, then you cannot have a desire for crème brulee, and you can’t engage in crème-brulee-directed behavior. So explaining your crème-brulee-directed behavior seems to require us to ascribe to you a concept of CRÈME BRULEE.

So concepts play a role in explaining behavior because they play a role in explaining possession of the propositional attitudes that themselves explain behavior. Concepts are also invoked to explain cognizers’ abilities to categorize the world. Concepts are taken to underwrite judgments about how things divide up into categories, and about conceptual and causal relations between different kinds of objects. Concepts, in this capacity, are devices for categorization; concepts allow their possessors to identify items as instances of those concepts, to relate those instances to instances of other categories, and to reason about and act appropriately toward things that fall under the concepts one possesses.

The semantic and categorizational features of concepts are related in the following way: the question of whether a person is correct or not in their categorization depends on whether or not the items in question fall within the extension of the concept in question. But it is the semantic properties of the concept that determine what falls within the extension of the concept. So the assessment of people’s behavior explained by appeal to the psychological properties of concepts depends upon the semantic features of concepts. For example, it is because it is crows that one’s CROW concept is about that

---

5 Not under that description, at least. Someone certainly could have a desire that is in fact for crème brulee without having a concept of it; for example, if that person has a desire for her sister’s favorite dessert, where her sister’s favorite dessert is in fact crème brulee. But that is a different desire (although in fact coextensive with the first).

6 For fuller discussion of these features of concepts, see Keil (1989), Murphy (2004), Prinz (2002).
explains why one’s CROW-involving judgments are true or false, and why one’s crow-directed behavior is appropriate or not. So the items that have psychological properties must also have semantic properties.

Recognition of these two separate sides of concepts’ explanatory role does not commit one to the explanatory priority of either in giving a theory of concepts. Fodor (2004), distinguishes between concept Cartesians, who hold that the fundamental role of concepts is to allow their possessors to think about their objects, and concept Pragmatists, who hold that the fundamental role of concepts is to allow their possessors to do certain kinds of things, such as to make inferences and discriminatory and categorizational judgments about their objects. But a complete theory of concepts should explain both how concepts have the semantic content that they do and how concepts are employed in generating behavior. This is not to say that there is no difference between these two roles, only that they are importantly connected. A theory of what concepts are should not ignore either feature.

1.2.3 The Publicity of Concepts

Any theory of concepts should explain how it is that the same concept can be possessed by a number of different individuals. Concepts are public entities, in the sense of being intersubjective. So a theory of the nature of concepts should explain how different individuals can be characterizable as having the same concept.

This has both a semantic and a psychological aspect as well. We need to be able to ascribe the same intentionally characterized states to different individuals in order to make sense of them as sharing propositional attitudes, and also in order to be able to offer
systematically related explanations of their behavior, to be able to explain how the behavior of different individuals resembles and differs. If we take the objects of propositional attitudes to be structured out of concepts, then we have reason to think of concepts as entities that can be shared by different individuals as well.\(^7\)

1.2.4 Concepts and Mental Representations

I have argued that concepts are mental representations. This is not to say that all mental representations are concepts. The possibility should be left open that some mental representations should not count as conceptual representations – that concepts might be a distinctive subset of mental representations. And in fact some such distinction seems presupposed by various different lines of thought in the cognitive sciences. For one example, Fodor’s (1983) version of modularity theory distinguishes between the mental representations that are involved in modular input systems and those that are involved in central processing. It seems reasonable to say that the representations that, on such a theory, are involved in the operations of central systems are a distinctive kind of mental representation; and it seems a reasonable policy to reserve “concept” to characterize this kind of mental representation. Similarly with theories that distinguish between procedural and declarative representations, or between perceptual and conceptual representations.\(^8\)

---

\(^7\) This requirement that concepts be public does not preclude there also being a sense in which people’s concepts may differ. It seems legitimate to speak in terms of my concept of x being different from your concept of x. The point about publicity is that we need to make space for the idea that two people’s concepts can differ in such a way that makes sense of how they can be different versions of the same thing.

\(^8\) I am not here endorsing any of these particular ways of making this distinction between conceptual and nonconceptual mental representations. Nor am I suggesting that all of them have the same distinction in mind.
This raises the question of what makes conceptual representation distinctive. Notice that this cannot be accounted for simply by appealing to the theoretical roles for concepts pointed out so far. As an example, consider Marr’s (1982) theory of vision. On at least some interpretations of Marr’s view, the primitive elements involved in constructing visual representations – for example, the representations of edges and lines – are themselves to be understood in representational terms. (Burge 1986) That is, they have a semantic value that they contribute to the construction of more complex visual representations. Similarly, such representations play a role in the explanation of behavior: it’s because people have visual representations with those contents that they react the way that they do under certain experimental conditions, for example. Finally, those representations are public in the sense required for generality of explanation; all normal human perceivers have visual systems that function in essentially the same way, and thus contain representations of edges and lines and the like. So, on this interpretation of Marr’s theory, the representations involved show the distinctive features of concepts that have been isolated so far. Yet a cognizer may not have a concept of a given property, even though that property is represented in her visual system.

But what conditions are distinctive of conceptual representation that would motivate ruling out, among others, low-level perceptual representations as concepts? Here are some intuitive considerations that might be mobilized in support of such a distinction:

- **flexibility**: concepts are representations that can be deployed in a wide variety of contexts, to a wide variety of ends and purposes, in a wide variety of projects. In
contrast, some conceptions of certain kinds of representations view them as constrained in their application to very restricted contexts. Some mental representations may have only domain-specific application. In contrast, conceptual representation is not limited in this way; it is a hallmark of a concept-using being that it can use the concepts that it has in a wide variety of situations.

- **access**: concepts are representations that are available to conscious thought and reflection, while other representations may not be. For example, low-level perceptual and linguistic representations are generally thought to be isolated from conscious access.

- **control**: concepts seem to be the kinds of things that are under the control of the organism, while other representations may not be. This is to say, roughly, that an organism can produce tokens of a conceptual representation more or less when it wants to. In contrast, certain other kinds of representations may be tokened only in the context of certain kinds of perceptual stimulation.

These considerations suggest that conceptual representations have distinctive features not shared by all mental representations. They are, however, less suggestive as to what the principled basis of such a distinction might be. I will return to this question in chapter 6, where I offer an account of what is distinctive of conceptual representation. For now I just want to make plausible the idea that concepts may be a distinctive subset of mental representations.

The foregoing considerations give us some idea of what kinds of things concepts are: concepts are intentionally characterized entities out of which the objects of...
propositional attitudes are constructed. They determine the satisfaction-conditions of the objects of the propositional attitudes by contributing their semantic value to those objects. They are also the elements which explain the behavior of the organism that possesses the concepts. But they are also distinguished from other sorts of mental representations by various distinctive features having to do with the ways in which they can be used. This gives us a rough characterization of the explanatory profile of concepts.

1.3 Concept Learning and the Argument for Radical Concept Nativism

With this preliminary characterization of concepts in hand, we are in a position to raise questions about how they are acquired. If concepts are items with features enabling them to satisfy the explanatory goals canvassed in the previous section, then possessing a concept must be a matter of bearing an appropriate relation to such an entity. The acquisition question, then, is a question about how a cognizer comes to satisfy the possession-conditions for a concept. The argument for radical concept nativism, begins by arguing that the explanation of how a cognizer comes to satisfy those possession-conditions cannot be a matter of learning.9

In presenting this argument it will help to introduce some terminology. Let a complex concept be one that contains other concepts as its constituents. Thus, TAPIOCA PUDDING is a complex concept, since it contains TAPIOCA and PUDDING as constituents. Let a primitive concept be one that is not complex. Let a lexical concept be a concept that is naturally expressible in English by a single word, and let a phrasal concept be a concept that is naturally expressible in English only by a multiple-word phrase. Finally,

9 Notice that it is a further question why the fact that concepts are not learned is supposed to entail that they are innate. I will address this aspect of the argument for radical concept nativism in chapter 2.
let a target concept be one whose acquisition is to be explained by a particular instance of a process of concept acquisition, and let a base concept be one possession of which is invoked to explain the acquisition of a particular target concept.\textsuperscript{10} 

1.3.1 Confirmation-Theoretic Models of Concept Acquisition

What does it take to learn a concept? Notice that concept learning is unlike at least some other kinds of learning, such as rote learning or memorization. This is because concepts have within their extension many more things than those we actually perceive in acquiring those concepts. A child may acquire BIRD from seeing just a few birds, yet the concept thus acquired has all birds within its extension. The most familiar model for such a process, where something general is produced on the basis of a finite number of particular instances, is inductive learning (Bruner et al. 1956; Fodor 1975). On this model, cognizers acquire concepts by making inductive inferences on the basis of the data they receive in experience. Specifically, acquiring a concept is a matter of forming and confirming hypotheses about the satisfaction-conditions of the concept. One has acquired a concept when one has accepted the correct hypothesis about what the satisfaction-conditions of that concept are. I will call such models Confirmation-Theoretic models of concept acquisition.

Confirmation-Theoretic models seem to provide the resources to solve two central problems for theories of concept acquisition. First, a theory of concept acquisition must solve the “doorknob/DOORKNOB problem” (Fodor 1998). That is, it must explain why perceiving things in the extension of a particular concept is conducive to acquiring that

\textsuperscript{10} The argument to follow derives in large part from Fodor’s work. I have introduced some of the terminology, but claim no originality for the actual arguments.
concept, rather than some other. Granted that DOORKNOB represents doorknobs, why is it that perceiving doorknobs leads you to acquire DOORKNOB, and not (for example) BUTTERSCOTCH or HUMMINGBIRD? After all, it is not always the case that, when someone acquire a concept through an experience of something, the concept that she acquire contains within its extension the thing that she has experienced. When a child acquires KANGAROO from a picture book, she acquires the concept of an animal, not (or at any rate not merely) of a picture of an animal, even though the experience she has is of a picture, not of the animal itself. So it is not trivial that the concepts we form from experience are so often of the very things that we do experience. What explains why we do?

Confirmation-Theoretic views provides a solution to the doorknob/DOORKNOB problem. According to Confirmation-Theoretic views, the reason that exposure to members of the extension of the concept gives us that concept (and not some other concept) is that we take members of the extension of a concept to be generally good sources of evidence for the right satisfaction-conditions of the concept within the extension of which they fall. We acquire SQUIRREL, for example, by perceiving squirrels, because squirrels are a good source of evidence about the kinds of features things have to have in order to fall within the extension of SQUIRREL. For instance, squirrels have four legs, bushy tails, eat nuts, and live in trees, all of which are features of things that fall within the extension of SQUIRREL. We do not, in contrast, acquire BUTTERSCOTCH from experiences of squirrels, because such experiences do not in general provide any information about the satisfaction-conditions of BUTTERSCOTCH.11

11 This explanation also has the virtue of explaining why KANGAROO can be acquired from pictures of kangaroos: because we take pictures of kangaroos to be sources of evidence for the satisfaction-conditions of KANGAROO.
Second, a theory of concept acquisition must solve the “qua-problem” (Sterelny 1989; Cowie 1999). That is, it must explain why one acquires the particular concepts that one does in perception, even though the things one acquires those concepts from fall within the extension of many other concepts as well. This particular thing may be a squirrel, but it is also many other things: a rodent, a mammal, a quadruped, a gray thing, a furry thing, etc. So why do cognizers acquire one concept rather than another by perceiving something that falls within the extension of many? *Qua* what do you acquire a general concept of the particular thing you perceive?

Confirmation-Theoretic views provide a solution to the *qua-problem* as well. On Confirmation-Theoretic views, the reason you acquire one concept, rather than another, through perceiving something that falls within the extension of both, is that you take what you perceive to provide evidence for what determines membership in the one category rather than the other. Which concept you acquire through your experiences depends on which hypothesis you are trying to confirm – that is, how you represent the hypothesis and the evidence that you take to bear on that hypothesis. So the reason you acquire SQUIRREL, and not MAMMAL, from perceptions of squirrels is that you take squirrels to provide a good source of evidence for hypotheses concerning the satisfaction-conditions of SQUIRREL, not of MAMMAL.

All this suggests that concept learning should be understood in Confirmation-Theoretic terms, according to which the kinds of psychological processes that underlie concept learning have to be processes that realize an inductive logic. These processes instantiate the formation and confirmation of hypotheses concerning the rule that determines the extension of the concept. Learning a concept is a matter of forming
hypotheses about what the concept does and does not apply to as a matter of general principle (i.e., learning the rule for applying the concept), and then testing those hypotheses against the data (i.e., the data about whether given instances are in fact instances of the concept or not). As a result of this process, she comes to learn the rule that governs which things the concept applies to and which not.

1.3.2 Learning and Circularity

Confirmation-Theoretic models of learning are committed to the claim that being able to form and test hypotheses is a necessary condition on anything’s being able to acquire concepts. This requires both that the cognizer be able to represent the hypotheses that it forms concerning the satisfaction-conditions of the target concept, and that the cognizer be able to represent the evidence that bears on the confirmation or disconfirmation of those hypotheses.

But this requirement threatens to generate an explanatory circularity. In order to acquire a target concept, the cognizer has to be able to form various representations already. This requires that the system already have the concepts necessary to form those representations. This means, at a minimum, that the cognizer must have the conceptual resources to represent the satisfaction conditions of the target concept in terms of some set of base concepts that it already possesses.

This implies that every primitive concept – every concept the satisfaction-conditions of which cannot be represented without that very concept – must be innate. YELLOW, for example, is primitive: it is impossible to say whether or not something satisfies the concept YELLOW other than by using YELLOW itself; i.e., by saying whether or not the thing in question is yellow. But if it is not possible to formulate the
satisfaction-conditions for a concept without using that very concept, then forming a hypothesis about the satisfaction-conditions of that concept cannot be a way of acquiring that concept, since you need to have the concept in order to form the hypothesis. But if the concept cannot be acquired by someone who does not already have it, then, it appears, one must have the concept to start with; that is to say, it must be part of one’s initial innate cognitive endowment.

So Confirmation-Theoretic models face an explanatory lacuna: they cannot account for the acquisition of *primitive* concepts, since primitive concepts are those whose satisfaction-conditions cannot be represented without employing that very concept. So no primitive concept can be learned. Rather, possession of primitive concepts must be presupposed in order to explain how complex concepts can be learned.

Thus we have the following argument:

1. Concepts are either acquired through learning or not.
2. If concepts are acquired through learning, they are acquired through hypothesis-formation and testing.
3. Primitive concepts cannot be acquired through hypothesis-formation and testing.
4. So: primitive concepts are not learned.

1.3.3 Conceptual Structure

At the beginning of the previous section, we saw that there was reason to want to accept Confirmation-Theoretic models of concept acquisition, since they promised to
explain how concepts could be acquired from interactions with things in the concept’s extension. Unfortunately, this model is not applicable to the acquisition of at least some concepts, namely the primitive concepts. The previous section argued for the conclusion that no primitive concept could be learned, on pain of circularity. But this circularity only threatens so long as the target concept being acquired is the same as the base concept. If the target concept could be acquired by invoking the mobilization of different base concepts, then the circularity would disappear.

It is, therefore, possible to respond to the argument of the previous section by denying its interest. One might grant that there must be some concepts for which the Confirmation-Theoretic model cannot provide an acquisition story, but hold that most concepts are not like that. Most concepts are acquired by mobilizing a fairly small and restricted base of primitive concepts. In particular, if most concepts are structured entities, composed in some way out of more basic concepts, then it is natural to suppose that acquiring a concept is a matter of constructing it from more primitive concepts.

The question then becomes whether it is reasonable to suppose that a relatively small number of primitive concepts are sufficient to account for the acquisition of the rest of our conceptual repertoire. In the terminology introduced above: are lexical concepts for the most part primitive concepts, or are most lexical concepts themselves complex constructions out of a significantly more restricted primitive conceptual base? If the latter is the case, then the Confirmation-Theoretic model of acquisition can be employed to explain the acquisition of at least the better part of our conceptual repertoire.

There are two kinds of views about how concepts might be acquired by constructing them from other concepts. On the first, a target concept is acquired by
assembling a *definition* of the concept from other concepts, which are either primitive concepts themselves or in turn are themselves defined in terms of other concepts. According to definitional views of conceptual structure, complex concepts are constructions out of more basic concepts that collectively encode necessary and sufficient conditions for something’s being in the extension of the complex concept. The base concepts collectively determine the satisfaction-conditions for the complex concept acquired on their basis.

Alternatively, the base concepts that are used to acquire the target concept might not provide a definition of the target concept in terms of the base concept, but might rather encode certain contingent but especially salient information about things in the extension of the concept. According to prototype theories of concepts (Rosch et al. 1976; Smith and Medin 1981), for example, concepts are structured entities composed of other concepts representing typical, salient, and diagnostic features of things in the extension of the concept. While such concepts don’t provide the satisfaction-conditions of the target concept, they encode information about what kinds of features things that do satisfy the concept tend to have. More generally, according to *nondefinitional* views of conceptual structure, complex concepts are constructions out of more basic concepts that don’t determine the satisfaction-conditions of the target concept. Acquiring a concept, on such views, is a matter of assembling it from concepts of the features that are typical of things that satisfy the target concept.

If a concept is either definitionally or nondefinitionally structured, the Confirmation-Theoretic model of concept acquisition can give a noncircular account of how that concept can be acquired. So, if a wide range of concepts are plausibly
structured in either of these ways, then the explanatory power of the Confirmation-Theoretic model can be salvaged. The next movement of the argument for radical concept nativism is to argue that most concepts are not plausibly structured in either of these ways. More precisely, the claim that the argument for radical concept nativism seeks to establish is that the vast majority of \textit{lexical} concepts – those concepts that are naturally expressed by a single term of a natural language – are unstructured.

The argument against concepts’ having definitional structure is simply that there do not appear to be definitions for any but obviously phrasal concepts. First, philosophical attempts at conceptual analysis have been largely unsuccessful. Philosophers have failed to give necessary and sufficient conditions for nearly every concept for which they have tried to give them, from the mundane (such as \textit{to paint}) to the momentous (\textit{knowledge, virtue}). Apart from certain toy examples, like \textbf{BACHELOR}\footnote{Lakoff (1987) points out that \textbf{BACHELOR} may not be a good example of a definitionally structured concept either. Potential counterexamples include priests (as individuals who satisfy the proposed definition but, intuitively, are not bachelors) and males from cultures who allow multiple marriages (as individuals who do not satisfy the proposed definition but, arguably, are bachelors). Bill Lycan points out that there are certain conventions that form lexical concepts with underlying definitional structure, such as terms for female animals: ‘doe’ means \textit{female deer}, ‘vixen’ means \textit{female fox}, etc. So there is certainly some evidence that some of our concepts have internal definitional structure. The arguments here are against the thesis that definitional structure is the general case.}, there appear to be few concepts that are not obviously phrasal but which turn out to be defined in terms of other concepts. And even if there turned out to be a few, that would be of little comfort to the proponent of definitionally structured concepts, since unless it turned out that most of our basic conceptual repertoire could be defined in terms of a relatively small number of primitive concepts, the Confirmation-Theoretic view could not claim to provide an account of concept acquisition that is at all general.
Second, there is little evidence to suggest that definitions have any psychological reality. That is, there is little evidence to suggest that people’s concepts have definitional structure. In particular, reaction time experiments suggest that the presence or absence of definitional representations does not seem to make a difference to psychological processing.\(^{13}\) This suggests that positing definitionally complex mental representations is psychologically unmotivated.

While definitional views of conceptual structure appear, as a matter of empirical fact, not to correctly characterize the mental states of cognizers, there is a good deal of evidence for the psychological reality of nondefinitional structures, such as the robust and widespread evidence for prototype effects on cognitive tasks. Despite this, nondefinitional views of conceptual structure appear to be unsatisfactory as well. This is because they fail to respect the requirement that concepts be compositional (see section 1.2.1 above).

Compositionality requires that the satisfaction-conditions of complex concepts be exhaustively determined by those of their parts. For if they were not, then simply having the simple concepts out of which the complex concept is composed would not suffice for generating the complex concept. But nondefinitional views of conceptual combination don’t respect this requirement. Prototypes, for example, often have emergent properties that none of the features out of which they are constructed have.\(^{14}\) So putting together a nondefinitional structure does not suffice for having the target concept, because the

---

\(^{13}\) See Fodor et al. (1980). As Fodor (1998:46) puts it: “It’s an iron law of cognitive science that, in experimental environments, definitions always behave exactly as if they weren’t there.”

\(^{14}\) The classic example of this is *pet fish*: the prototypical pet fish is about 3 inches long and lives in a fishbowl, but neither the prototypical pet nor the prototypical fish have these features.
concepts that are supposed to constitute the structure do not collectively entail the right application-conditions for the concept. So nondefinitional views of concept combination cannot explain the compositionality of concepts.

If most lexical concepts are neither definitionally nor nondefinitionally structured, then they must not be structured at all. So, almost all lexical concepts are primitive concepts. If this is right, then almost no lexical concept can be learned, since almost no lexical concepts can be noncircularly acquired Confirmation-Theoretically.

Thus we have the full form of the argument for radical concept nativism:

1. Concepts are either acquired through learning or not.
2. If a concept is acquired through learning, it is acquired through hypothesis-formation and testing.
3. If a concept is acquired through hypothesis-formation and testing, then it is structured.
4. If a concept is structured, it is either definitionally or nondefinitionally structured.
5. Lexical concepts are not definitionally structured.
6. Lexical concepts are not nondefinitionally structured.
7. So: lexical concepts are not structured.
8. So: lexical concepts are not acquired through hypothesis-formation and testing.
9. So: lexical concepts are not learned.

The conclusion of this argument, in turn, has been supposed to imply that lexical concepts are innate. The force of this argument, then, is that we have an innate
conceptual system at least as expressively powerful as any natural language we could learn to speak. In particular, we do not significantly increase the expressive power of our conceptual system by learning new concepts.

Radical concept nativism is, at first sight, wildly implausible. How could concepts such as LEGUME, CHIFFON, ALTERNATOR, or GASOLINE possibly be innate? But this depends on what the significance of claiming that a concept is innate in fact is. I will return to these issues in chapter 2. In the remainder of this chapter I will explain the central thesis on which the argument turns.

1.4 The Conceptual Mediation Thesis and the Representational Necessitation Thesis

The way I have presented the argument for radical concept nativism emphasizes the question of whether or not concepts are structured as the central issue, and this has historically been the feature of the argument that has commanded the most attention. But in fact the question of conceptual structure is inessential to motivating the problem for theories of acquisition.

To see this, note that we can separate two questions about concept acquisition. One question is about the *processes* of acquisition; the other is about the *products* of those processes. The Confirmation-Theoretic model of acquisition is an answer to the first question; it is an attempt to explain what is distinctive about processes that eventuate in a cognizer’s possession of new concepts. The claims about conceptual structure, on

---

15 To take a recent example, Margolis and Laurence (2002) explicitly identify conceptual structure as the crucial issue in the argument for radical concept nativism. I will discuss their view in more detail in chapter 4.
the other hand, are claims about the products of concept acquisition; they are claims about features of the concept thus acquired.

But thinking of concept acquisition in Confirmation-Theoretic terms does not mandate thinking of complex concepts as literally constructed out of other concepts, of containing more basic concepts as parts. What the view requires is just that a cognizer acquire new concepts by mobilizing concepts that she already possesses; it does not require further that the concept thus acquired contains those concepts as literal parts.

Failure to appreciate this may be due in part to a confusion of the conventions for naming concepts with those of giving a structural description of a concept. Consider the concept GREEN OVAL. If “GREEN OVAL” is a structural description of that concept, then clearly the concept includes GREEN and OVAL as constituent concepts. But if “GREEN OVAL” is simply a name for a concept, then the fact that we use a two-word phrase to refer to that concept does not entail that the concept itself is a complex mental representation. It need not be a complex representation even if its extension is exactly those things that are both green and oval-shaped. For we can introduce a new symbol into our representational system – for example, ‘Ω’ – that has exactly the same extension as GREEN AND OVAL-SHAPED, but which does not contain those concepts as parts. ‘Ω’ is a new primitive symbol, even though adding it to a system does not increase the expressive power of the representational system to which it is added. Similarly, ‘Ω’ could be acquired by a cognizer in the way that the Confirmation-Theoretic model specifies, so long as that cognizer already possessed the concepts GREEN and OVAL.

So the idea that concepts are acquired by mobilizing concepts one already possesses is not wedded to the idea that the concepts thus acquired are themselves
composites of other concepts. That a process of acquisition requires the use of some concepts in order to acquire others does not in turn require that the product of that process be a structured entity.\textsuperscript{16}

So – despite the orientation of the bulk of the literature – the argument for radical concept nativism does not depend on any controversial thesis about whether or not most concepts have internal structure. On what then does it depend? I believe the argument depends on a particular thesis about what it would take for a concept to be learned. I will call this presupposition the Conceptual Mediation Thesis (CMT):

\begin{enumerate}
\item[(CMT)] Necessarily, all acquisition of new concepts by a cognitive system is mediated by concepts that the cognitive system already possesses.
\end{enumerate}

Commitment to CMT underpins the Confirmation-Theoretic learning model. The idea that concepts are acquired by forming and testing hypotheses concerning the satisfaction-conditions of the target concept reflects the idea that one needs to use concepts in order to acquire other concepts. But the Confirmation-Theoretic learning model reflects just one particular way of being committed to CMT.\textsuperscript{17} The general idea is that learning a concept, whatever exactly the processes by which this occurs, is a matter of using concepts that you already have to acquire new concepts. This reflects a more general thought about what is distinctive of learning: that learning is a matter of acquiring

\begin{itemize}
\item\textsuperscript{16} Lakoff (1987) makes a similar point about prototype theory in warning against the inference from the existence of prototype effects to the conclusion that concepts have prototype structure.
\item\textsuperscript{17} In chapter 4 I will consider other accounts of concept acquisition that accept CMT without accepting the standard Confirmation-Theoretic view.
\end{itemize}
new beliefs. If this is an accurate characterization of learning generally, then concept learning specifically must be a matter of acquiring new beliefs. But acquiring new beliefs presupposes that you have the capacity to have those beliefs, and that in turn requires that you have the concepts needed to form those beliefs. So concept learning requires that you already have the conceptual resources necessary to acquire the beliefs constitutive of acquiring a new concept.

The fact that Confirmation-Theoretic views are committed to CMT explains why arguments against Confirmation-Theoretic views have so generally been taken to yield nativist conclusions. If CMT expresses a constraint on learning model of concept acquisition, then any concept that is not acquired in accordance with CMT is therefore unlearned. If one supposes that concepts are either learned or innate, then the failure of learning models to explain acquisition leads to the conclusion that concepts must in general be innate.

I have identified CMT as the central thesis acceptance of which motivates arguments for radical concept nativism. My goal in the major part of this work will be to argue that CMT is false. But nothing in my argument will depend on just how concepts are distinguished from other mental representations. Nor will it depend on any specific claims about the kind of mediation involved in acquiring new concepts. So I will argue against CMT by arguing against the much weaker Representational Necessitation Thesis (RNT):

18 In saying this I am not claiming that there is no difference between concepts and other kinds of mental representations. My claim here is that my arguments against CMT will not presuppose any such distinction. I will return to these issues in chapter 6, where I will argue for a distinction between concepts and other kinds of mental representations on the basis of the features of flexibility, access, and control, and discuss how to understand the way in which concepts one does possess can mediate the acquisition of new concepts.
(RNT) Necessarily, if a cognitive system acquires any concept, then it already possessed some mental representations prior to acquiring that concept.

RNT makes no claims about the role that these mental representations play in acquiring new concepts. In fact, for all that RNT says, these mental representations could play no role at all in acquiring new concepts. Nor does RNT require that any mediation be specifically conceptual mediation; RNT would be true if a cognitive system needed some mental representations that were not concepts in order to acquire new concepts. But if RNT is false, then CMT is false as well; if it is not necessary for a cognitive system to possess any other mental representations in order to acquire concepts, then it is not necessary that mental representations play any mediating role in acquiring concepts.

So my arguments, in what follows, should be understood as directed against RNT. I will argue that cognizers can acquire some concepts without having any mental representations at all. Despite this thrust of my argument, however, it will often be convenient to phrase my discussion in terms of CMT rather than RNT in the following chapters. This is because the explanatory goals of a theory of concept acquisition often lead theorists to hold views that are best characterized by CMT rather than simply by RNT. But I want to stress that, when I argue against CMT, I will not do so on the basis of some particular understanding of what a concept is, or of precisely how concepts mediate the acquisition of new concepts. My arguments against CMT are meant to apply to RNT as well.
1.5 Conclusion

So there appear to be three possible positions for explaining concept acquisition:

(a) We could accept radical concept nativism.

(b) We could accept CMT/RNT and argue that this acceptance is compatible with explaining concept acquisition.

(c) We could reject CMT/RNT as a constraint on an account of concept acquisition.

In my view, neither (a) nor (b) is promising. On the one hand, attempts to explain concept acquisition consistent with accepting CMT face problems of the kind canvassed above. On the other hand, thinking of concepts as innate seems untenable as well. That suggests we should explore option (c).

But option (c) can seem problematic as well; indeed, it can seem not to be an option at all. For it has been thought that any view of concept acquisition that doesn’t accept CMT is *ipso facto* nativist. I’ll explore the reasons for this in chapter 2, and motivate the idea that an explanatory theory of concept acquisition need not require appeals to prior concept possession.

Even given that it is possible, however, there are reasons for concern about the viability of option (c). For it is not as though acceptance of CMT is unmotivated. On the contrary, as I explained above, acceptance of CMT derives plausibility from various explanatory virtues of learning models such as the Confirmation-Theoretic model. Any theory of concept acquisition that proposes to dispense with CMT needs to show how
those explanatory goals can be met without appeal to CMT. In chapter 4 I will discuss more recent views that attempt to avoid nativist conclusions without rejecting CMT. I will argue that these views are unsatisfactory, but that there are important lessons to be learned by considering them. That will pave the way in chapters 5 for an alternative theory of concept acquisition; one that explains how concepts can be acquired without commitment to CMT.
2.1 Introduction

In chapter 1 I presented the argument for radical concept nativism, and explored the challenge that this argument poses to giving a theory of concept acquisition. According to this argument, learning models of concept acquisition explain acquisition of concepts by appeal to the forming and testing of hypotheses concerning the application-conditions of the concept. Such models have the virtue of being able to explain why concepts are acquired from encounters with instances of them. However, these models of acquisition appear to be unable to explain the acquisition of primitive concepts, and there is good reason to think that concepts are not in general structured. Hence, concepts cannot be learned.

The most common understanding of this argument is as arguing for the conclusion that concepts are innate. This is a point on which nativists and nonnativists about concepts agree - that the argument I presented in the previous chapter is an argument for radical concept nativism.

But while this claim sounds radical, it is also somewhat obscure, because it is not clear what commitments are carried by claims that psychological structures are innate. One commitment that clearly is not intended is that there is no such thing as conceptual development at all. As Richard Samuels (2002) notes, there is a “baseline sense” of acquisition, according to which some psychological structure C is acquired by a subject S.
just in case there is some time \( t \) prior to which \( S \) lacked \( C \) and after which \( S \) possessed \( C \).

In this sense of acquisition, it is uncontroversial that concepts are acquired. So concept nativism should not be committed to denying that there is *any* sense in which concepts are acquired.

Furthermore, it is uncontroversial that the character of our conceptual repertoire is responsive to the nature of our interactions with the environment. So concept nativism should not be committed to denying that environmental interactions play a role in determining our mature conceptual system. What is controversial is the form a theory of such interactions should take. So concept nativism should not be opposed to the thesis that there is such a thing as conceptual development. Instead, concept nativism should be treated as a particular proposal for how to understand cognitive development.

Acknowledgement of these points, however, makes it less clear what the debate between nativists and nonnativists concerns. In the opening sections of this chapter, I will argue for a view of the content of nativist claims that has been articulated and defended by Samuels and Fiona Cowie. According to this view, claims about the innateness of psychological structures are claims about what kinds of explanations are appropriate for the presence of those structures. Specifically, claims about the innateness of psychological structures are claims that the acquisition of those structures are not amenable to psychological explanation. According to this way of understanding the nativist claim, debates over whether or not concepts are innate are debates about whether or not psychology has the resources to explain concept acquisition.

I argue that this conception of nativist claims explains why the conclusion of the arguments from the previous chapter – that concepts are unlearned – has been taken to
argue for radical concept nativism. If concepts are unlearned, and learning is the distinctively psychological mechanism for acquiring concepts, then concept acquisition is not a candidate for psychological explanation.

I then turn to the question of why learning should be thought to be the only candidate for giving a psychological explanation of concept acquisition. I argue that this is in virtue of learning models’ commitment to the Conceptual Mediation Thesis, and is motivated by acceptance of the Computational Theory of Mind. I then argue that computational explanation should not be taken to constrain psychological explanation more generally.

2.2 Methodological Nativism

Nativism is generally presented in the first instance as a negative thesis. In the arguments from chapter 1, for example, nativism is inferred from the failure of Confirmation-Theoretic models to explain concept acquisition. But this raises the question of what distinctive claim is being made in characterizing a psychological structure as innate.  

Samuels (2002) argues for a view of the commitments of nativist claims that he calls primitivism. Cowie (1999) defends a similar position. I will refer to this view of the import of nativist claims as a methodological conception of nativism. According to

---

19 In fact, failure to find a satisfactory answer to this question has led to some skepticism about the usefulness of innateness as a theoretical construct; see Oyama (1985), Griffiths (2002).

20 Cowie’s view may be more radical; she seems at times to suggest that nativism is committed, not just to the impossibility of psychological explanation of concept acquisition, but to the impossibility of any scientifically respectable explanation of concept acquisition. But this seems too strong; presumably nativists are not debarred from giving, for example, an evolutionary explanation for the presence of concepts in the initial state. In what follows I will focus on Samuels’ version of the methodological conception of nativism.
this view, claims about the innateness of some psychological structure are claims about
the viability or otherwise of certain kinds of explanation for the presence of the structure
in question.\textsuperscript{21} More specifically: according to Samuels, the nativist claim about concepts
is that the environmental interactions that eventuate in the possession of a concept cannot
be explained in psychological terms. To claim that a psychological structure is innate, on
this view, is to claim that its presence is not something that psychology has the resources
to explain. As Samuels puts the claim, a psychological structure\textsuperscript{22} is innate just in case it
is a \textit{psychological primitive},\textsuperscript{23} where “a psychological structure S … is a psychological
primitive just in case:

1. S is a structure posited by some correct scientific psychological theory.

2. There is no correct scientific psychological theory that explains the acquisition
of S. (Samuels 2002: 246)

Notice the claim is not that a psychological structure is innate if no \textit{current}
psychological theory can explain its acquisition. Rather, it is that, for an innate

\textsuperscript{21} I am not claiming that this \textit{exhausts} the content of nativist claims. This is not presented as an analysis of
what “innate” means. All I claim is that the lack of psychological explanation is one commitment of
nativist views, and that it is a commitment of all nativist views. Notice too that this reading of nativist
claims is compatible with taking “innate” to be something like a natural-kind term (see Prinz 2002).
Perhaps genetics will prove to vindicate claims of innateness by showing how to account for the presence
of some psychological structures. But claims about the innateness of cognitive structures need not, just as
such, be committed to any particular view of the appropriate kind of explanation for the presence of such
structures. Nativists can disagree about what kind of explanation would vindicate nativist claims.

\textsuperscript{22} Such as a concept, a belief or other propositional attitude, a learning mechanism, a perceptual or
cognitive module, for example. A complete inventory of psychological structures presumably must await
the resources of a completed psychology. I will return to this point momentarily.

\textsuperscript{23} “Primitive” in this sense should not be confused with the notion of a primitive concept introduced in
chapter 1. The latter sense of “primitive” contrasts with “structured”, while the former contrasts with
“acquired via psychologically explicable means”. The argument for radical concept nativism purports to
show that all primitive concepts are psychological primitives, but that is the conclusion of the argument,
not something to be assumed at the outset.
psychological structure, there is no scientific psychological explanation to be had of its acquisition. Notice too that this is not to claim that the acquisition of that structure is inexplicable, period. There may be biological or physiological explanations of the presence of such structures. What is ruled out is just that the presence of such structures could be explained by appeal to other psychological structures and processes.

On this view, the conclusion that concepts are innate does follow from the arguments presented in chapter 1. The conclusion of those arguments was that concepts cannot be learned. Learning is there conceived to be the distinctive candidate psychological process to explain concept acquisition. Since the primitivist thesis just is that what cannot be explained in psychological terms is innate, the claim that concepts are innate is simply another way of putting the claim that concept acquisition is not psychologically explicable. Thus the methodological conception of nativism explains the inference from “unlearned” to “innate”.

Samuels motivates the methodological conception of nativism by arguing that it makes sense of the debates between nativists and nonnativists about various psychological structures. According to the methodological conception, debates about whether some psychological structure is innate or not are debates over what kinds of explanation are appropriate for how a cognizer acquires (in the “baseline sense” of acquisition) that structure. Nativists hold that the explanatory resources of psychology are insufficient to account for the acquisition of such structures, while nonnativists hold that psychology can explain how cognizers acquire them. In the case of concepts, this understanding of nativist claims and nativist-nonnativist debates explains why the arguments from the previous chapter should be understood as arguments for nativism. So
the methodological conception of nativism appears at least to capture a necessary condition for something to be innate.

However, problems have been raised for this way of understanding nativist claims. Wilson (2004) objects to this methodological conception of nativism that appealing to the notion of a completed scientific psychology raises objectionable epistemic problems for evaluating nativist claims. We have little idea at present what a completed and correct scientific psychological theory might look like, and what its domain and explanatory power might be. So we have little idea at present what sorts of entities and processes it might take as primitive. So, if this were the right way to understand nativist claims, then we would have little idea how to evaluate nativist claims in psychology. This suggests that the methodological conception of nativism is inadequate.

This objection can be answered. It is true that, if our psychological theories are incorrect, then the claims we make that depend on their truth will likely be incorrect as well. So there is always a danger of fallibility with claims that rely on theories that are potentially false. But that does not prevent us from understanding claims that particular psychological structures are innate. Furthermore, we can at least evaluate them provisionally: if, given our current psychological theories, we can give a plausible explanation for the acquisition of a psychological structure, then we have as good a reason to deny that structure is innate as we do to accept our current theory. On the other hand, if we cannot give an acquisition story within the constraints of the theory we accept, then we have reason to believe that the structure may be innate. Either way, our lack of a complete and correct psychological theory does not prevent us from
understanding claims made with reference to it. And this is all that is required to explain what debates between nativists and nonnativists concern.

A greater worry is that the methodological conception of nativism is too liberal in what it counts as innate. According to the methodological conception, a psychological structure is innate if its presence cannot be explained by psychology. But there are various kinds of cases where a psychological structure apparently results from a non-psychological process, but where it is not plausible to claim that the structure is innate. Suppose, for example, that one were to somehow acquire a concept or a belief as a result of a blow to the head. The explanation for how one acquired that concept or belief would not, it seems, make reference to a psychological process, but would rather appeal to the underlying neurological processes caused by the impact that realize the belief or concept thus acquired. In a similar vein, Fodor (1975) imagines a Latin pill, ingestion of which gives one full knowledge of Latin by reorganizing one’s underlying brain chemistry. In such a case there might be no psychological explanation of how the pill works; nevertheless, the knowledge the pill confers is not innate in the person who takes it. Less fancifully, there are cases where brain lesions and diseases appear to have the effect of generating novel cognitive structures. In none of these cases does it seem plausible to maintain that the relevant structures are innate. But the methodological conception of innateness seems required to count these as cases of innate psychological structures, since by hypothesis their acquisition is not explicable in psychological terms. This suggests that the methodological conception of nativism does not capture the right contrast between nativists and nonnativists.
Samuels proposes to deal with such cases by adding a “normalcy” qualification to the primitivist account: a cognitive structure is innate in a cognitive system only if that cognitive system would acquire that cognitive structure in the normal course of events. While Samuels does not attempt to spell out this condition in any detail, he suggests that it is the right kind of strategy to solve this problem. This raises the question whether there is a plausible understanding of “the normal course of events” that will rule out cases such as those mentioned above, where a psychological structure is not plausibly innate but neither is it acquired by psychologically explicable means.

Appealing to a “normal course of events” is in fact another popular strategy for giving an account of the commitments of nativist claims. André Ariew (1996, 1999) argues that innateness is a matter of the degree to which a trait is canalized, where the degree to which a trait is canalized is a matter of whether it would emerge in a range of different developmental environments. The intuitive motivation for this view is that, while innate traits require some environmental stimulation for their development, they do not require any particularly specific kind of environmental stimulation. For example, in the case of physiology, human beings will develop secondary sex characteristics so long as they survive into puberty, and for this they require input from their environment to whatever extent is needed to keep them alive and developing normally. But the development of secondary sex characteristics does not require any particularly specific environmental stimulation; lots of very different developmental environments are conducive to humans developing secondary sex characteristics. Similarly, perhaps, to say that a psychological structure is innate is to say that any of a wide range of developmental environments would be conducive to developing that psychological structure.
But this does not seem to be the kind of understanding of “normal course of events” that the nativist about concepts needs. This is because the concept nativist holds that many of our particular concepts are innate. But it is not plausibly true that particular concepts are acquired simply in virtue of having some normal cognitive development. While a wide range of developmental environments are all compatible with acquiring some concepts or others, which particular concepts one acquires depends on the specific character of one’s interactions with the environment. But a normal cognitive development does not as such require any particular set of environmental interactions.

I pointed out at the beginning of this chapter that nativism should not be construed so as to deny that the character of our mature conceptual repertoire is responsive to the specific character of our experience. This suggests that trying to characterize the content of concept nativist claims by a general appeal to “normal development” is misguided. A proper characterization of concept nativist claims should be such as to make them compatible with the idea that very specific kinds of developmental interactions might be required to acquire particular concepts. On the other hand, as we have seen, the construal of nativism as a purely negative thesis about the limitations of psychological explanation is unable to draw the contrast between what is innate and what is acquired through adventitious nonpsychological means. What the nativist needs is an account of what is distinctive about the processes that, on the nativist account, give rise to concepts, and an explanation of why those processes count as the development of innate concepts rather than the acquisition of new concepts.

It is plausible that part of what nativism about psychological structures entails is that there is no explaining how those structures are acquired in psychological terms.
Reflection on the excessive liberalism in what such accounts count as innate, however, suggests that this cannot be all there is to nativist claims. What nativists need is an account of what it is about the development of a certain psychological structure that marks it as an innately specified structure. This requires a distinctive positive view about the process of acquisition. Appeal to the “normal course of development” is insufficient, since that sort of appeal fails to answer the question why it is that such development proceeds. The nativist needs a story about what distinctive about the process of development of those psychological structures that she takes to be innate, and this story needs to rule out cases of acquisition by brain injury and the like as cases of accidental acquisition.\footnote{This is not the nativist’s burden in particular; a psychological theory of concept acquisition needs an account of what distinguishes psychologically explicable acquisition from accidental acquisition as well. But this burden is easier to meet for the nonnativist, since nonnativist views are not plausibly understood in purely negative terms.}

In the next chapter I will examine the most developed nativist account of this process; the idea that the development of innate concepts is a matter of having that concept \textit{triggered} by experience. In the remainder of this chapter I will address the question of why, in light of the argument presented in chapter 1, one might be tempted to conclude that concepts are innate. What, in other words, makes the argument presented in chapter 1 an argument for radical concept \textit{nativism}?

\subsection*{2.3 Nativism, Computationalism, and Psychological Explanation}

I have argued that a core commitment of nativist claims is what I have called the methodological conception of nativism. According to the methodological conception of nativism, the claim that a given cognitive structure is innate carries with it a commitment
to the lack of a psychological explanation for the presence of that cognitive structure. I think that this claim makes sense of the inference from the conclusion of the arguments presented in chapter 1 – that concepts cannot be learned – to radical concept nativism. The thesis that concepts are innate, according to the methodological conception of nativism, follows from those arguments together with the thesis that learning is the distinctively psychological mechanism by which concepts are acquired. This gives us the following argument:

(1) Concepts are not learned.

(2) Learning is the distinctive psychological mechanism by which concepts are acquired.

(3) If the acquisition of a concept is not explicable by psychological mechanisms, then that concept is innate.

(4) The acquisition of most lexical concepts is not explicable by psychological mechanisms. (1, 2)

(5) So: most lexical concepts are innate. (3, 4)

Establishing (1) was the burden of the arguments from chapter 1, and (3) expresses the methodological conception of nativism set out in the earlier sections of the present chapter.

So the question, in evaluating this argument, is whether there is reason to accept (2). Why think that any properly psychological explanation of concept acquisition must be in terms of a learning model of concept acquisition?
As I noted at the end of chapter 1, what is distinctive of learning models of concept acquisition is their commitment to what I called the Conceptual Mediation Thesis. That is, learning models hold that the acquisition of new concepts must be mediated by concepts one already possesses. The claim that learning models are the distinctive psychological mechanisms for concept acquisition is motivated by the thought that psychologically explicable acquisition requires the mobilization of concepts one already has. That is, CMT is supposed to express a constraint on a distinctively psychological theory of concept acquisition.

But why should this be? What motivates treating CMT as a condition on a psychological theory of concept acquisition? I think that the reason doing so arises from acceptance of the Computational Theory of Mind.\(^{25}\)

The Computational Theory of Mind is a theory of the nature of the underlying states and processes that implement paradigmatic psychological phenomena such as reasoning. According to computationalism, psychological processes are realized by computational processes. The mind, on this view, is fundamentally a kind of computer.

If concept acquisition is to be a psychological phenomenon, then, according to the computationalist account of what psychological phenomena fundamentally are, concept acquisition must be implemented by computational processes. But computational processes are defined over representational states. So, if concept acquisition is a computational process, then it must itself essentially involve representational states, and those representational states must be in place before the processes that are defined over

\(^{25}\) I think this motivation for CMT may at least in part underlie Fodor's (1975) discussion of concept acquisition. I have found no explicit statement of the motivation for this position, however, so the discussion to follow is my own reconstruction.
them can operate. But, if, in order to acquire concepts via psychologically explicable means, you already need to have some representational states, then, ultimately, you will need to have some representational states that are not acquired via computational processes. But, given the methodological conception of nativism presented in the previous section, this is tantamount to saying that those representational states are innate. Alternatively, suppose there is reason to think that concepts are not acquired by computational processes. But, according to computationalism, psychological processes are computational processes. So, evidence that concepts are not acquired by computational processes is evidence that concepts are not acquired by psychological processes, and again, given the methodological conception of nativism previously articulated, that is evidence that those concepts are innate. So acceptance of the Computational Theory of Mind argues in favor of accepting CMT as a condition on what counts as a psychological explanation of concept acquisition, in a way that explains why a psychological explanation of concept acquisition is the alternative to concept nativism.

These considerations provide us with an explanation of the inference from “unlearned” to “innate” that connects the failure of learning models for concepts with the thesis of radical concept nativism. The question now is whether this inference is compelling. I will now argue that it is not. The fact that concept acquisition is not explainable in accordance with CMT, in a way that avoids commitment to radical concept nativism, should not be taken to show that concept acquisition is not psychologically explicable at all. Concept acquisition need not be computationally explicable in order to be psychologically explicable.
In arguing this, however, I will not argue against the thesis that many psychological processes are computational processes, nor will I argue for an alternative conception of the underlying architecture of cognition. Instead, I will argue that computational processes are not the only psychologically explicable processes there are.

I begin with a distinction between three explanatory levels. (Marr 1982; Sterelny 1990; McClamrock 1995) The first level is the level of specifying the cognitive capacities in question. This level has been variously called the computational level, the ecological level, and the task level. This level of explanation specifies what the cognitive system does and what it has the ability to do.

The second level is the level of specifying the processes by which the system performs those capacities isolated at the first level of explanation. This level has been variously called the algorithm level, the computational level, and the process level. This level of explanation specifies how the system is able to do what according to the first level of explanation, it has the ability or competence to do.

The third level is the level of specifying the physical realization of the processes isolated at the second level of explanation. This level is commonly called the implementation level. This level of explanation specifies how the system is built, in virtue of which it is able to do the things it does in the particular ways it does them.

The Computational Theory of Mind, as a theory of the processes subserving psychological processes, is a theory at the second level of explanation. That is, the thesis that psychological processes are computational processes is a thesis about the nature of the processes that allow the system to perform the tasks that it can perform.
But that leaves open the question of which tasks those are. Before we can offer a computational theory of the processes which subserve a particular capacity, we must have some understanding of what that capacity or set of capacities is. That is, we require a theory at the first level, a theory of the nature of the cognitive capacities that a system has.

This means that an important part of the task of psychology is to give an account of the kinds of tasks and capacities that cognitive systems have. Providing first-level theories is an important part of the enterprise of scientific psychology. Some of the tasks and capacities thus isolated at this level will have interesting and significant commonalities between them, in ways that suggest similarities in how the system in question performs them. Where we find such commonalities, there we have reason to think that those will be reflected in similarities of underlying processes at the second level of explanation.

The methodological distinction I am pointing to here is similar to one that Dennett (1981) draws, between what he calls *intentional systems theory* and *sub-personal cognitive psychology*. The former is a theoretical system the aim of which is to provide a catalog of the tasks and capacities of cognitive systems, while remaining neutral on the internal structures in virtue of which the cognitive system is able to perform those tasks. The latter explanatory task is the burden of sub-personal cognitive psychology, which is

---

*Although Dennett is also concerned there with the relation between the pretheoretic categories of folk psychology and the post-theoretic categories of scientific psychology. While this is not an issue on which I take a stand here, it is worth noting that, since the formulation of level-one theories is itself part of scientific psychology, the same issues regarding the legitimacy of the categories used in such a theory arise in the case of scientific psychology just as much as they do in the case of folk psychology.*
devoted to the question of the structural features of the system in virtue of which it can perform those tasks.

I have argued for a distinction between level-one theories and level-two theories. I have also argued that psychology is concerned with theories at both levels. Computationalism, as I pointed out earlier, is a level-two theory; it is a theory about how a system performs the tasks that it performs. Since computationalism is a level-two theory, and since psychology is concerned with providing level-one theories as well, there is space for psychological theorizing that is officially noncommittal about the processes that instantiate cognitive tasks.

So psychological theory-building does not as such require a commitment to computational processes. I will now argue that there are positive reasons to think that at least some psychological processes are not subserved by computational processes, and that this fact is compatible both with computationalism being a correct theory of cognitive architecture and with those processes being unproblematic examples – indeed, paradigm examples – of psychological processes.

My argument begins with the multiple realizability of cognitive systems. (Putnam 1967; Fodor 1974) There are various kinds of physical structures that could potentially implement cognitive structures. Different configurations of physical states can be sufficient for the same psychological states. In the terms introduced above, there is no one-one mapping between first-level and third-level theories. Nor, and for much the same reasons, need there be any one-one mapping between second-level and third-level theories. The same computation can be performed by different kinds of physical systems. Nor again need there be any one-one mapping between first-level and second-level
theories. The same task can be performed by different computational structures. This relative independence of explanatory levels allows investigations at one level to constrain investigations at other levels without fully determining them.

So much may be relatively uncontroversial. What may be more controversial is the idea that a natural grouping of phenomena at one level may have very different kinds of underlying explanations at a lower level. Consider the various level-one tasks and capacities ascribed to cognizers by psychology, such as the capacity to have propositional attitudes of various kinds, to have emotional states, to consciously access one’s own psychological states, to perceive their environment, to reason and plan, and so on. There is no reason to doubt that these capacities are underwritten by very different level-two processes. But computationalism claims that the phenomena unified under the heading of psychology will receive a unified explanation at the underlying level of the cognitive architecture. That is, although different psychological phenomena will certainly receive different kinds of computational explanations, still something’s status as a psychological phenomenon at all will be vindicated by there being some computational process that underwrites it. I want to suggest that this may not be true either.

The central argument for computationalism is its ability to model reasoning by construing inference as a species of computation. (Turing 1950; Fodor 1975, 1987) But while reasoning is a central cognitive process, and a paradigm psychological phenomenon, it is by no means the only one.27 So the fact that reasoning is a likely

27 Computational models are also popular in the study of perception, though there are disputes over how psychologically realistic these models are. I will consider such models in more detail in chapter 5. Interestingly, Fodor (2000) argues that computational models are more plausible as models of relatively self-contained perceptual systems than of central cognitive systems. I do not find Fodor’s arguments for those claims persuasive, but I cannot go into those issues here. As I will argue in chapter 5, I think the features of perceptual systems that Fodor claims make them ideal candidates for computational modeling
candidate for explanation by appeal to underlying computational processes argues for the computational approach to psychological phenomena more generally only insofar as there is independent reason to think that such phenomena in general involve reasoning. Conversely, to the extent that we think reasoning is not involved in some psychological phenomenon, we should be suspicious of the ability of computational explanations to model the phenomenon adequately.

So the question, with respect to concept acquisition, is whether there is reason to think that reasoning is involved in the processes that lead to the acquiring of concepts. But such a reason for denying that concept acquisition requires such vindication is close to hand. In fact it has already been provided, in the arguments from chapter 1 that show that learning models of concept acquisition are untenable. The fact that learning models are supposed to be underwritten by computational models, and that computational models are bad models of concept acquisition, gives us reason to think that computational processes do not underwrite the paradigmatically psychological phenomenon of concept acquisition.

This turns Fodor’s argument on its head. Rather than supposing that the failure of computational explanations for concept acquisition shows that concepts must be innate, we can take it to show that computational models are not the right models to explain concept acquisition. But to suppose that this latter claim is itself a reason to regard concepts as innate is to suppose that computational models are the only legitimate level-two processes capable of legitimizing level-one psychological phenomena, and we have seen reason to resist that claim in general anyway.

\[\text{(specifically, their purported modular character) actually make them bad candidates for computational systems.}\]

53
This is not to claim that computationalism is false as a theory of the processes that underwrite some paradigmatically cognitive processes; I am not here taking issue with computationalism as an adequate theory of cognitive architecture. But I am claiming that, where there is reason to suppose that a computational explanation for a paradigmatically cognitive process is not forthcoming, that may be a reason for questioning the appropriateness of requiring a computational explanation instead of ruling the phenomena out of psychology’s explanatory domain. The fact that some psychological phenomena cannot be given a computational explanation is no threat to the adequacy of computational explanations of other psychological phenomena.

I am not insisting, a priori, that concepts simply cannot be innate. But I am claiming that the innateness of concepts cannot be established by showing that there is no computational explanation for their acquisition. Whether the lack of a viable computational explanation for a given phenomenon shows that the phenomenon is not psychologically explicable depends on whether we have independent reason to think that it should receive a computational explanation. But in the case of concept acquisition, it seems to me that we have positive reason to think that it should not. Computational explanations are paradigmatically explanations of the processes that underwrite the cognitive capacities of a mature cognizer. The fact that these explanatory tools don’t apply well to explaining how cognizers come to be possessed of those capacities does not show that computationalism is false nor that developmental psychology is not a kind of psychology. What is shows is that not all facets of psychology depend on being vindicated by the same underlying cognitive architecture.
These considerations would be less compelling if computationalism were indeed, as Fodor and others have claimed, “the only game in town.” But I have argued here that there is no reason to think that that’s the case. In Chapter 5 I will present an account of concept acquisition that rejects CMT, and I will argue that it nevertheless keeps concept acquisition within the explanatory purview of psychology. I have tried to argue here that it would be premature to think that such a view was ipso facto a nativist view of conceptual development. The fact that learning theories of concept acquisition have traditionally been Confirmation-Theoretic theories, and that concept acquisition has proved hard to explain by the lights of those models, may well indicate a deficiency in those models. But, just for that reason, we should be wary of assuming those those exhaust the resources available to psychology in general.

All this is, I think, as it should be. Claims about the innateness of concepts or any other cognitive structure should be justified on the basis of plausible proposals for how those structures are acquired. Some of these proposals will appeal heavily to the cognizer’s interactions with the environment and the peculiarities of her individual ontogenetic history, while others will emphasize maturational factors and the role of the cognizer’s phylogenetic history. In evaluating these proposals, we should not be swayed by claims about what style of explanation must be correct, but should look instead to what style of explanation seems to yield most promising results.

In this section I have argued that computational explanation does not exhaust the resources of psychological explanation more generally. To the extent that acceptance of CMT is implicitly supported by this presupposition, as I suggested above that it was, these arguments remove that reason for accepting CMT.
2.4 Conclusion

In chapter 1 I set out the argument for the conclusion that concepts are unlearned. My project in this chapter has been to investigate why that conclusion has been taken to argue for radical concept nativism. I think that we can now see why that is so. According to the Computational Theory of Mind, psychological processes are computational processes. Thus, if a process is not a computational process, then it is not psychologically explicable. But to claim that some cognitive structure is innate is, at least in part, to claim that it is not psychologically explicable. So, the argument goes, unless the processes of concept acquisition are computational processes, then concepts are innate. But Confirmation-Theoretic models exhaust the options for computational explanations of concept acquisition. Therefore, if Confirmation-Theoretic models are untenable, then concepts are innate.

I have also indicated how to resist this line of argument. Computational explanations need not underwrite all legitimate psychological explanations, not even all legitimate psychological process explanations. This is consistent with accepting that computationalism gives by and large a correct account of the cognitive architecture of mature cognizers, or at least of central features thereof.

What is left now is to evaluate proposals for how concepts are acquired. In the next two chapters I will consider two different kinds of views. In chapter 3 I will examine Fodor’s nativist view of conceptual development. In chapter 4 I will look at recent views that accept CMT and try to avoid the conclusions of the arguments from chapter 1. I will argue that neither of these strategies are successful. But I have been
concerned to argue here that neither do they exhaust the options. Seeing how they fail will put us in a position to construct a more adequate theory, one that shares with Fodor’s view a rejection of traditional learning models, but one that nevertheless exhibits concept acquisition as a paradigmatically psychological phenomenon.
CHAPTER 3  NATIVIST VIEWS OF CONCEPTUAL DEVELOPMENT

3.1 Introduction

According to the methodological conception of nativism examined in chapter 2, disputes over innateness or otherwise of psychological structures are disputes over what kinds of explanations for the presence of those structures are likely to be appropriate. But nativism with respect to a certain kind of cognitive structure cannot be established simply by arguing against the availability of a psychological explanation for that structure. For the lack of a psychological explanation for a certain structure is compatible with that structure having been acquired through some kind of accident. This highlights two requirements on a nativist account of conceptual development. First, such an account needs to explain how concepts are acquired in the “baseline sense” of acquisition explained in chapter 2. The primary issue between nativists and nonnativists is not whether concepts are acquired or not, because there is a sense of acquisition in which it is uncontroversial that concepts are acquired. Second, such an account must explain what makes the acquisition, in this sense, of some class of concepts count as the development of innate concepts, rather than the acquisition of concepts by accidental means. So a nativist account of conceptual development needs a theory of the processes that result in a cognizer’s having the psychological structures that she does, and an explanation of what is distinctive about those processes that means that the concepts a cognizer develops in consequence of having them should count as innate.
In this chapter I will critically discuss the most well-known nativist account of conceptual development in the literature: Jerry Fodor’s (1981) view that concepts are triggered by experience rather than learned from experience. I will argue that Fodor’s view faces various difficulties accounting for various facets of concept acquisition. While I do not claim that these problems are insurmountable, I do think that they suggest that we should revisit the issue of whether concept acquisition is amenable to psychological explanation. I argued in the last chapter that the arguments for radical concept nativism may depend on an impoverished conception of the resources available to psychological explanation; I will argue that the problems a nativist account of conceptual development faces give us additional reason to expand our conception of the resources psychology has for explaining the origin of our concepts.

3.2 The Triggering Conception of Concept Acquisition

According to the triggering conception of concept acquisition, concepts are triggered by experience in essentially the way in which, according to ethology, some behaviors are triggered by experience. Ethological examples of triggering include the following behavior that a duckling engages in toward the first moving object that it sees after opening its eyes, or the aggressive behavior of a male stickleback when presented with a red patch. (Tinbergen 1952; Lorenz 1970) In such cases, a complex behavioral response is released by a stimulus that is informationally impoverished relative to the behavior that it elicits. When it sees the moving object, the duckling engages in behavior that is an appropriate response to its mother. But the duckling does not learn that the first moving object it sees is its mother; at any rate it does not learn it from the stimulus that
elicits the behavior. Nor is it plausibly true that its following behavior is evoked by a
judgment that the first moving thing it sees is likely to be its mother. Instead, the
behavior is probably an innately specified response to this particular stimulus. Similarly,
the male stickleback does not learn that red patches are indicative of other male
sticklebacks; at any rate it does not learn it from the stimulus that elicits the behavior.
Nor is its aggressive behavior plausibly evoked by its judgment that the red spot it sees is
likely to signal a potential rival. As in the case of the duckling, the male stickleback’s
behavior is probably an innately specified response to this particular stimulus. The fact
that the stimulus is relatively simple, and the elicited behavior relatively complex, are a
reason for regarding the behavioral reaction as innate rather than acquired.

Fodor argues that, much as in the case of triggered behavior concepts are elicited
by fairly impoverished stimuli rather than being learned. Learning is, in Fodor’s
terminology, a “rational-causal” process, while triggering is a “brute-causal” process. To
say that a process is “rational-causal” is essentially to say that it is a Confirmation-
Theoretic process in the sense of chapter 1. To say that a process is “brute-causal”,
conversely, is to deny that it is such a process.

That a process of concept formation is brute-causal entails, at least, that the
experience that triggers the concept need stand in no straightforward evidential relation to
the concept that it triggers. Fodor emphasizes that the relation between a concept and its
triggering stimulus can be arbitrary, in that the latter need provide no evidence about the
satisfaction-conditions of the concept that it triggers. Just as the duckling imprints of the
first moving thing it sees, despite the fact that motion is not universally good evidence of
motherhood, so, Fodor claims, having a triggering experience is sufficient to acquire a
concept whether or not that experience provides good evidence about the concept that one thus acquires.\textsuperscript{28}

Furthermore, whether or not the triggering experience does in fact provide good evidence about the satisfaction-conditions of the concept that it triggers, according to Fodor the cognizer who acquires the concept on the basis of having the triggering experience does not acquire that concept in virtue of \textit{appreciating} any evidential relation that might obtain between experience and concept. The cognizer does not engage in any process of \textit{reasoning} from the content of the experience to the content of the concept. Instead, merely having the triggering experience is causally sufficient for acquiring the concept.

So, in contrast to the Confirmation-Theoretic model of concept acquisition, the triggering model holds that (a) the relations between the content of a concept and the experiences that give rise to that concept can be arbitrary, and in particular the latter need not provide good evidence for the former, and (b) having the relevant triggering experience is all that is required for acquiring a concept; in particular, no processes of reasoning on the part of the cognizer are involved. This latter claim is what contrasts most directly with the Confirmation-Theoretic view of concept acquisition, and it seems to be what leads Fodor to argue that his is a nativist view. According to the previous

\textsuperscript{28} Sterelny (1989) argues that the relation between stimulus and concept can’t be completely arbitrary, but must be constrained by evolutionary considerations, because what is innate must, according to Sterelny, be explicable as a result of evolutionary processes. Similarly, the relation between the duckling’s seeing a moving object and imprinting on it is not completely arbitrary, because there are good evolutionary reasons why this experience should trigger this behavior in the duckling (namely, that in the duckling’s environment, the first moving object that the duckling sees is overwhelmingly likely to be its mother. Fodor would likely not be swayed by these considerations, since he expresses pessimism about evolutionary explanations. However, it is open to him to reformulate the distinction between learning models and triggering models in terms of the degree of allowable arbitrariness in the relation between stimulus and concept.
chapter, acceptance of the Computational Theory of Mind provides a motivation for taking Confirmation-Theoretic views to exhaust the resources for explaining concept acquisition in psychologically respectable terms. This would explain why rejection of the Confirmation-Theoretic view of concept acquisition would be seen as tantamount to an acceptance of nativism: since triggering is not a Confirmation-Theoretic process, it is therefore not a psychologically explicable process. Therefore triggering views are nativist views.

I have already argued against this motivation for nativism. The fact that some process is not a process of reasoning – and more generally, that it does not involve already-possessed representations – does not mean that it is not a psychological process. So the fact that a view of concept acquisition is modeled on the ethological notion of triggering does not obviously entail that the view is committed to innate concepts. This will be important in chapter 5, for my own view of how concepts are acquired also rejects the idea that concept acquisition is a matter of operations on mental representations. But, I argue, it is nevertheless a psychological theory of how concepts are acquired. But for present purposes I will put this issue aside. In the remainder of this chapter I will argue that the triggering account of concept acquisition faces problems even on the assumption that it is a genuinely nativist view of conceptual development.

3.3 Problems With Triggering

The brief sketch of Fodor’s triggering view in the previous section raises three questions:
(1) What is being claimed to be innate prior to the activation of a concept by a triggering experience?

(2) What is a triggering experience?

(3) How does the triggering view account for the fact that people so often do acquire concepts from encounters with instances of them?

I think that the triggering account of concept acquisition can answer none of these questions satisfactorily.

3.3.1 What is Innate?

Jerry Samet and Owen Flanagan (1989), Fiona Cowie (1999), and Prinz (2002) point out that there are two ways to read the claim that concepts are innate prior to being activated by encounters with their triggering stimuli. On the stronger reading, the claim is that concepts are present in the mind prior to activation in just the way that they are after being activated. That is, the mind is furnished with a stock of mental particulars that are independently identifiable prior to activation; the only difference between these and activated concepts is that the former are not available for use by the cognizer herself. Cowie calls these untriggered concepts “protoconcepts”.

On the weaker reading, the process of triggering does not activate a preexisting mental structure. Rather, to claim that concepts are triggered is just to claim that a cognitive system will form certain concepts in response to certain inputs. There need be nothing independently identifiable as a mental object prior to triggering. On this reading, the claim that triggering “activates” a concept is simply amounts to the claim that a certain experience leads us to form a particular concept. This claim carries with it no
commitment as to how this process works, or in virtue of what it is that this experience gives rise to this concept.

So, the stronger reading posits innate mental objects that become concepts when activated by the appropriate triggering experience. Prior to triggering, these protoconcepts are not available to the cognitive faculties of the cognizer. The weaker reading, by contrast, is noncommittal about exactly what is innate prior to triggering; the claim is simply that triggering experiences lead to the formation of concepts. The difference between the two views is that the stronger requires some independently intelligible way to identify protoconcepts prior to triggering, while the latter does not.

The problem with the weaker reading of the claim that concepts are triggered by experience is that it is too weak to mark out a distinctively nativist view of concept acquisition. That is because the claim that experience leads to the formation of concepts – even to claim that experience leads to the formation of concepts in the way that it does because of features of our cognitive architecture – is fully compatible with a learning model of concept acquisition. As I pointed out above, any model of concept acquisition, whether nativist or nonnativist, holds that the concepts we form is sensitive to the particular character of our experience. What is at issue is the nature of the processes by which experiences of a particular kind lead us to form the concepts that we do.

So the weaker reading of the claim about protoconcepts does not draw a contrast between nativist and nonnativist accounts of conceptual development. This is a problem because, as we have seen, concept nativism cannot simply be defined by exclusion; the nativist requires an account of what is distinctive of innate concepts. But the weak thesis about protoconcepts does not provide such an account.
These considerations suggest that the stronger reading is required to give an account of what is distinctively “brute-causal” about the relationship between concept and input. And this in turn requires some account of the individuating properties of protoconcepts. Cowie (1999) considers two possible ways of accounting for these properties of protoconcepts: that protoconcepts are mental objects with formal/syntactic properties but no semantic properties, and that protoconcepts are mental objects with narrow content but no wide content.

Recall from chapter 1 that concepts are semantically evaluable items with causal powers. Thus they have two kinds of properties: those in virtue of which they have their satisfaction-conditions, and those in virtue of which they enter into mental processes. According to the Computational Theory of Mind, the features of concepts that allow them to enter into mental processes are to be understood as syntactic features over which computational processes are defined. Roughly, syntactic features are the intrinsic or local features of a symbol; something like the “shape” of the symbol. Computational processes are sensitive to the syntactic features of a symbol; it is because the symbol has the shape that it does that it is treated in the way that it is by the processes which operate on it.

So one way of understanding the notion of a protoconcept is as a mental object having only syntactic properties. On this view, a protoconcept acquires its semantic value in consequence of a triggering experience. Acquiring a concept is a matter of having a preexisting symbol endowed with a semantic content in consequence of undergoing the relevant triggering experience. What is innate, on this view, is the syntactically structured symbol. Therefore innate protoconcepts can be individuated prior to activation by reference to their syntactic properties.
Cowie objects to this proposal that “[i]t seems implausible to think that the mind at birth has all these uninterpreted symbols actually stored … in memory.” (Cowie 1999: 81) But it is not obvious that this is any more implausible than the thesis of radical concept nativism itself, nor that it is implausible in any different way. But concept nativists generally reject plausibility judgments about what is innate as reflecting a refusal to engage with philosophical argument or empirical considerations. Presumably a radical concept nativist would say the same thing about the claim that innate syntactically structured symbols are implausible. If accepting innate syntactically structured but uninterpreted symbols is what explaining conceptual development requires, then we have reason to accept them, intuitive implausibility aside.

But I think there is a further problem with the proposal that protoconcepts have syntactic but no semantic properties. Recall that, for according to computationalism, the syntactic properties of a concept are those in virtue of which it enters into the causal processes that realize mental operations. So it appears that to grant that a protoconcept has such syntactic properties is to grant that it is already available to the cognizer for use in mental operations. For if a protoconcept has its syntactic properties prior to triggering, then it already has all the features it needs to be used in mental operations. But if this is right, then it is no longer clear in what sense a triggering experience is supposed to make the concept available to the system, since it seems that the protoconcept is already able to be used by the system. It is true that, prior to triggering, the concept does not have a semantic value, so operations with that concept have the character of operations with a purely formal symbol. But, on the computationalist view, whether or not the symbol in
fact represents anything makes no difference to how the operations are actually carried out.

The triggering theorist might respond that while the protoconcept has its syntactic properties prior to triggering, the triggering experience is required in order for those properties to figure in those causal processes that are constitutive of cognitive activity. While the protoconcept already has its syntactic properties, on this view, something in the system prevents the protoconcept from interacting with other cognitive structures. The role of a triggering experience here is not to confer any new property to the protoconcept, but is simply to release it from whatever prevents it from being used by the system.

Certainly this is a coherent position for the triggering theorist to hold. However, I think it strains credulity in ways beyond the initial implausibility of radical concept nativism. This view requires us to suppose not only that everything required for any concept to enter into cognitive processes is innate, but additionally that there are cognitive barricades that prevent most of those concepts from actually being used in cognitive processes until some external conditions are met. But why, if all that is needed for any given concept to be used is already innate in the system, should that system be positively prevented from using them? Why should a cognizer have abilities that it is positively restricted from using? It seems that the triggering theorist owes us an argument for why cognizers develop in this way.29

The second possibility that Cowie considers that protoconcepts have narrow content prior to being triggered, and acquire a wide content in consequence of being

29 While I think that evolutionary explanations should constrain attributions of innate structures, I do not mean to be relying on that thought here. I am simply suggesting that building a wide range of abilities into a given system, and then additionally building in preventions on the exercise of those abilities, would be a strange and counterproductive way to build a system.
triggered. Therefore, protoconcepts can be individuated by their narrow contents. Understanding this proposal requires a brief explanation of narrow content.

Externalists about mental content hold that the intentional properties of a mental state do not supervene just on facts about the cognizer’s internal states. (Putnam 1975; Burge 1979) The belief I would express with the words “Water is wet” is made true by the facts about H₂O, while the belief that my twin-earth doppelganger would express in the same way is made true by the facts about XYZ, even though our internal states may be completely indistinguishable. But noting that our mental states have different semantic properties leaves the question of how to characterize what I and my doppelganger have in common. In response, some theorists have introduced a notion of narrow content, which is conceived to be a kind of content that I and my doppelganger share.

There are two standard views of narrow content in the literature: either as a function that takes contexts as input and delivers wide contents as output, or as a kind of conceptual role. On the first view, the narrow content of a mental state determines the wide content of a mental state from the context in which the concept is tokened. My WATER concept refers to H₂O because the relevant context is Earth, while my doppelganger’s concept refers to XYZ because the relevant context is Twin-Earth. What we share is the function that determines those references from the contexts we are in. On the second view, the narrow content of a mental state is the conceptual role that it has: the causal/inferential relations that it bears to other concepts, to perceptual inputs, and to

---

30 There may also be a third way of understanding narrow content, simply as that which explains the intrinsic causal powers of mental representations. This may be what Fodor (1987) understands by narrow content. To the extent that narrow contents taxonomize by causal powers, they seem to me to raise the same issues as taxonomy by syntax does, so I will not discuss that option further here.
behavioral outputs. While my doppelganger’s water concept may have a different reference from mine, we still have the same sensations when we perceive the stuffs that our water concepts refer to in our respective contexts, and behave in the same ways toward them.31

The question then is whether either of these notions of narrow content can provide plausible candidates for the individuating properties of protoconcepts prior to triggering. Cowie argues that neither is a viable option. Against the idea of protoconcepts as having narrow conceptual roles, she argues that conceptual roles are learned, not innate. Against the idea that protoconcepts are functions from contexts to wide contents, she argues that it reduces simply to the claim that concepts are acquired from their instances; in other words, this proposal reduces to the weaker reading of innate protoconcepts presented above.

It is not clear that a concept nativist should be moved by either of these considerations. There is no obvious reason why conceptual roles could not be innate. However, this position may face the same problem I raised above for the idea that the syntax of protoconcepts is innately prespecified. If narrow conceptual roles are innate, and if having a narrow conceptual role is all that is required for the cognizer to be able to use the concept, then it is unclear in what sense the concept is not available to the cognizer prior to triggering. In which case, the proposal fails to account for the baseline.

---

31 This possibility requires that conceptual roles be themselves narrowly specified, and there’s controversy about whether this is the right way to understand conceptual role. It is possible to combine these two conceptions of narrow content, to have a view according to which the (narrow) conceptual role of a concept figures into determining its wide content. Block (1986) has such a view, which I will look at in detail in chapter 4. I should emphasize here that I am not endorsing either of these notions of narrow content, nor am I endorsing the notion of narrow content at all.
sense of acquisition that is the common explanatory target of all accounts of concept acquisition.

Note that this position cannot avail itself of the response that the proponent of innate syntactic features has to this problem, namely that a representation can have syntactic properties and still be prevented from participating in cognitive processes. Syntactic features are supposed to be intrinsic properties of a representation – properties that do not depend on any relations the representation bears to other things. For that reason, it is coherent to maintain that a representation has its syntactic properties even though it is prevented in some way from participating in cognitive processes. Conceptual roles, on the other hand, are relational properties of a representation – they depend on the relations the concept bears to other representations. So there is no way to make a distinction between the conceptual role of a representation and its availability for use by the system. So, if a representation’s conceptual role is innate, then everything necessary for that representation to participate in cognitive processes is thereby innate.

Cowie’s second criticism of identifying protoconcepts by their narrow content is that it simply reduces to the claim that concepts are acquired from their instances. But it is not clear that this is so. First, it is unclear why acceptance of narrow content in this sense is a presupposition of any theory of concept acquisition. But, if the first claim indeed reduces to the second, then one would expect that acceptance of narrow content in this sense would be no more controversial than accepting the claim that concepts are indeed acquired from their instances. But acceptance of narrow content seems in fact to be a more substantive commitment than simply the acceptance that there is such a thing as concept acquisition.
However, narrow content in this function sense may not be a promising candidate for explaining protoconcepts either. According to this view of protoconcepts, what is innate is a function from the context of one’s tokening a concept to the wide content of that concept. But, if this is innately specified, then presumably it operates whenever the relevant concept is tokened. But if that is right, then it is not clear what is acquired when a new concept is triggered. Suppose that, prior to triggering, one has a mental structure that returns the wide content \textit{CAT} as output when given one’s current context as input. Then, every tokening of this mental structure will count as a tokening of a \textit{CAT} concept. But then claiming that this function is innate seems to sufficient for claiming the concept is innate.

One might respond that, since the features of context that the function takes as input are not innate, the wide content of the concept is not innate either. Since, for example, whether the function is deployed on Earth or twin-Earth determines whether one’s concept refers to cats or twin-cats, and whether one is on Earth or twin-Earth is not itself innately specified, whether one has a \textit{CAT} or a \textit{TWIN-CAT} concept is not innate either. This preserves a role for experience in explaining the acquisition (in the baseline sense) of concepts.

The first thing to say about this proposal is that, even if it is successful, it provides an even more limited role for experience than the initial triggering proposal. On this view, experience determines whether one has a \textit{CAT} or a \textit{TWIN-CAT} concept, but does not determine whether one has a concept of one of these at all. That is to say, it is still innately predetermined that one has a concept with the narrow content \textit{these catlike}
the role of the triggering experience is just to determine to which of the underlying kinds that has the superficial features of cats the concept in fact refers. Relatedly, this proposal would only preserve a role for experience in determining the content of those concepts that have the relevant kind of underlying essence. Any concept the reference of which does not shift with context in this way will still have its wide content innately.

Furthermore, I think this proposal faces a dilemma. Suppose an individual has an innate protoconcept with the narrow content (in the function sense) of \textit{catlike thing}, and now suppose she has an experience of a cat. If she were then transported to twin-Earth, would her concept refer to twin-cats? If it would, then it would seem that nothing relevant has been changed by the triggering experience. Prior to triggering, she had an innate function from contexts to contents, and now she has the same function. So the content of her concept has not been changed by the triggering experience, which seems to mean that the triggering experience has been irrelevant. On the other hand, suppose that, following the triggering experience, her concept is such that it only refers to cats, regardless of the context she is in. Then what is it about the experience that is supposed to guarantee this, since, by hypothesis, the triggering experience is of something indistinguishable from twin-cats?\footnote{I use this as a neutral expression to capture what cats and twin-cats have in common.}

\footnote{Note that my point does not depend on assuming that, if someone is to have a concept of something, they must be able to distinguish it from other things. The triggering theorist holds that causal interaction with instances is sufficient to acquire the relevant concept. On the present construal of that view, this requires that contextual features themselves play a causal role in fixing the wide content of the concept that is generated by the function in the course of having the triggering experience. The only candidate contextual features for playing such a causal role would seem to be features of the object of the triggering experience. So the question the triggering theorist has to answer is: which features are those? They cannot be the superficial features of the experienced object, since by hypothesis those are what cats and twin-cats have in common. But there do not seem to be any other plausible candidates.}
A final possibility, one that Cowie does not discuss, is that the wide content of concepts may be innate. But I do not think this proposal will work either. Content, according to contemporary naturalistic theories of content, is at least constituted by some sort of mind-world causal relations. But untriggered concepts are not supposed to be able to stand in causal relations before they are triggered. If so, then untriggered concepts cannot have wide content.

We have been looking for some account of the individuating properties of protoconcepts that would allow us to say what is innate prior to triggering. Any such account must both characterize the innate contribution to concept acquisition in a way that marks out a distinctively nativist view of the origin of concepts, and account for the baseline sense of acquisition that is the common explanatory target of all accounts of conceptual development. I have suggested that the particular proposals available to the triggering theorist have problems satisfying both of these desiderata.

In his (2001) response to Cowie, Fodor offers another potential option for explaining the innateness of concepts. Fodor claims that an externalist about the content of concepts can still coherently hold that concepts are innate, because she can hold that “the semantic properties of ‘protoconcepts’ supervene on their *dispositions* to enter into causal mind-to-world relationships.” (2001: 137) Fodor’s proposal here seems to be that protoconcepts can be individuated by their triggers.34 In other words, perhaps the problem with the previous proposals was that they attempted to identify the innate contribution of protoconcepts by features they have *prior* to triggering. But perhaps

34 Fodor continues: “Maybe what makes a mental representation a token of the protoconcept type CAT is its disposition to be triggered by cats.” (2001: 137)
instead the way to spell out a distinctively nativist view of conceptual development is by reference to what would cause a concept to be made available to a cognitive system.

So maybe questions about what is innate prior to triggering can be answered by specifying the triggers themselves. But I think that it is doubtful that the notion of a triggering stimulus can do the work that Fodor requires of it.

3.3.2 Problems with Triggers

According to Fodor, triggering views of concepts are distinctive in holding that the experiences that make a concept available to a cognizer are both necessary and sufficient to occasion acquisition of the concepts of which they are triggers. This contrasts with Confirmation-Theoretic views of concept acquisition, according to which acquiring a concept requires, in addition to having the relevant experiences, mobilizing concepts one already has in the formulation and testing of hypotheses that result in acquisition of the target concept.

This difference suggests that the way to mark out the difference between nativist and nonnativist views is, not by reference to the properties of some structure that is independently identifiable prior to its being triggered, but rather by reference to the triggering experiences that would yield acquisition of a given concept. The proposal, then, is that what is innate is a disposition or set of dispositions a cognizer has to form a particular concept in virtue of having one of a particular set of experiences.

In order for this proposal to work, there be some way to categorize triggering experiences for a particular concept such that having one of them is necessary, and that having any one of them is sufficient, to acquire the concept for which these are triggering
experiences. The question that needs to be answered, then, is: what is it about a set of experiences that makes them all triggers for the same concept?

The most natural thought is that these proprietary triggers are just the intentional objects themselves; you acquire the concept DOG from perceptual contact with dogs, paradigmatically. What it is about a triggering experience that makes it a triggering experience for DOG is that it is an experience that has dog as part of its intentional content. But, prima facie, this move is unavailable to the triggering view. First, there is a worry about circularity. For an experience to represent to you a dog, you plausibly have to already have the concept DOG. If so, then appeals to the intentional content of an experience can’t explain how you acquire the concept, because the fact that the experience has the intentional content that it does presupposes that one possesses the concept that the having of the experience was supposed to explain.

Alternatively, suppose that your experiences can contain dog as part of their intentional content even if you lack the concept DOG. Even so, not all experiences that result in acquiring a concept plausibly have the intentional object of that concept as part of their intentional content. This is because of something that Fodor himself stresses, namely that the links between the concept acquired and the experiences that lead to it can be arbitrary. The major difference between the triggering view and learning views was supposed to be that the triggering view does not require that your experience provide evidence for the satisfaction-conditions of concepts, in order to form concepts from them. Instead, the triggering experience itself is supposed to suffice for activating the relevant concept. So the triggering view seems committed to holding that the intentional content of an experience is irrelevant to its role in triggering a concept.
If the intentional content of an experience is not relevant to its role as a triggering experience, then is there any other common feature of all experiences the having of which is sufficient to acquire a given concept? The prospects for identifying such a common feature do not look promising. There seems to be no plausible candidate for a nonrepresentational feature of an experience that all triggering experiences for a given concept might share. If experiences from different modalities can trigger the same concept, then there will be no physical property of the stimulus that will be a viable candidate.

The triggering theorist might respond by emphasizing the arbitrariness of the experience-content relations. The fact that certain experiences trigger a particular concept is just a brute fact about human cognition, and there is nothing that triggering experiences must share in virtue of which they are triggers for this concept. It is simply a fact, not susceptible of further explanation, that the triggers for a given concept are what they are.35

This yields the picture of the innate contribution to concept acquisition as rather like a giant look-up table. When a cognizer undergoes a triggering experience, which concept she acquires in consequence is determined by which concept that experience is associated with. What is innately prespecified, on this view, are the associations of triggering experiences with concepts.

This story would need to be significantly complicated to deal with the role of context in determining whether a given experience triggers a given concept. Whether a given experience or set of experiences leads one to form a concept can, it seems, depend

35 Waiving, for the moment, doubts about the extent to which the relation between experience and concept can be arbitrary – see note 1.
on the situation in which one has the relevant experience. It is a strength of the Confirmation-Theoretic view that it can account for the role of context in terms of its effect on what hypotheses the cognizer forms and tests. In order to accommodate this feature of concept acquisition, the triggering view would seem to need to have all possible effects of context on the relation between triggering experiences and triggered concepts built in as well. While this is not impossible, it does add a significant additional layer of complexity to the triggering account.

Furthermore, the view of triggering experiences as innately prespecified faces problems with the plasticity of triggering experiences. What experiences give rise to a given concept does not seem to be fixed. This is most evident with the development of various forms of public representational media. For example, it is possible for an individual to acquire a concept of, as it might be, penguins even though they had never seen a live penguin before, through reading books about penguins, seeing pictures of penguins, having penguins described to them, etc. In fact, uses of public representational media would seem to be a fairly common way for people to acquire concepts. But such ways of acquiring concepts are only made possible by the introduction of new representational resources into the concept-learners environment.

It would even seem to be possible to acquire a concept of a kind without direct perceptual contact with any of its instances and also without relying on public representations of them. Suppose that you encounter a distinctive kind of animal footprint in the woods. By investigating further, you find further distinctive signs: spoor, claw markings on trees, etc. (Suppose also that you are familiar with the distinctive signs of all the other kinds of animals in the forest, so that you have good evidence that all
these signs are from the same kind of animal.) It seems plausible that these signs could lead you to acquire a concept of the animal in question, even though you’ve never been in direct perceptual contact with any of them, and even though you’ve never encountered any public representations of them.

Now, it could be replied that the relations between every experience that could in principle be had by a cognizer, and the concept that would be formed as a result, are innately fixed. But this seems to stretch the bounds of plausibility. The triggering view was to be offered as an explication of the thesis that concepts are innate consistent with offering a viable model of the processes that lead to concept acquisition, in the “baseline” sense” that is the common explanatory target of all theories of concept acquisition. But, if it is held that the relations between concepts and triggering experiences is simply a brute inexplicable fact, then the triggering view does not provide an explanation of those processes, but simply notes that there are such processes.

3.3.3 Problems with Arbitrariness and Invariance

It is distinctive of triggering views of concept acquisition, as opposed to learning views, that the relation between concept and triggering experience can be arbitrary. This is because, according to triggering views, the relation between concept and triggering stimulus is “brute-causal”, and this means at least that there is, or at least need be, no evidential relation between the environmental input and the concept formed as a result. For example, in the case of the fledgling duck, what seems to lead the duck to regard a particular object as its mother is its perception of a moving object immediately after
hatching. It is not the case that the duck has an experience that it takes as evidence for
the existence of a kind satisfying the concept MOTHER.36

But the fact that the concept-trigger relation can be arbitrary raises a question
about why then, as a matter of empirical fact, concepts are so often acquired from
interactions with their instances. Since Confirmation-Theoretic models posit a evidential
relation between experience and concept in virtue of which one can take the experience
as evidence for the satisfaction-conditions of the concept, and since encounters with
instances plausibly are good sources of evidence about the satisfaction-conditions of the
concept, learning models predict that cognitive systems should, by and large, acquire
concepts from interaction with instances of them. Triggering models, on the other hand,
do not explain why this should be true.37

The triggering view also has problems explaining the observed course of
conceptual development. It has been observed that children tend to acquire concepts of
middle-sized object first, with concepts of actions, of properties, and more abstract
concepts coming later in development. (Rosch 1976; Carey 1978) But this broad
constancy in conceptual development is potentially at odds with the triggering view.
Since, according to a triggering view, all that is required for the acquisition of a concept
is the relevant triggering experience, the triggering view apparently should predict that
children acquire all these concepts whenever they are able to acquire any concepts at all.

36 According to Fodor, at any rate. But the story is somewhat more complicated. Given a choice of moving
objects on which to imprint, ducks will preferentially imprint to the object that looks and moves most like
an adult duck. (Staddon 1985) While this does not show that ducks in fact are rational-causal learners, it
does seem to be evidence of more complicated processing than the triggering view recognizes. In
particular, it is not clear that such processing does not involve mediation by other of the duck’s
psychological states.

37 This is just the “doorknob/DOORKNOB” problem discussed in chapter 1.
Fodor’s response to this problem is to invoke a hierarchy of triggers. The idea is that not all concepts are such as to be triggered by experience regardless of the psychological makeup of the cognitive system. Perhaps, rather, some concepts can only be triggered once other concepts have been acquired. So, for example, perhaps you can only acquire TABBY once you’ve acquired CAT. Similarly, perhaps you can only acquire concepts of properties of things, like RED or SWEET, once you’ve acquired some concepts of things that could potentially have those properties.

But this proposal begins to place necessary conditions on concept acquisition that include other psychological states, in particular the presence of other concepts. So to make this move is to claim that the acquisition of at least some concepts is mediated by the possession of other concepts. No longer is the having of a triggering experience always sufficient to activate a concept; whether or not it is sufficient depends on what other concepts you already possess.

It is true that acceptance of a hierarchy of triggers needn’t commit one to the requirement that the experiences provide evidence that the concept-learner uses to form new concepts. But that view at least offers an explanation of why it is that the acquisition of some concepts should be mediated by the possession of other concepts. In the absence of some explanation for why this should be so, the hierarchy proposal appears to be an ad hoc addition to the triggering view, one not in the spirit with the motivating concerns of the view, and one that adds no explanatory power to the view, but which is introduced simply to allow the view to accommodate the available data.

Certainly none of the problems that I have raised in this section show that the triggering view of concept acquisition is incoherent, nor that it is irremediably flawed.
But perhaps enough has been said to show that the major nativist alternative to Confirmation-Theoretic models of concept acquisition faces explanatory problems as well. Fodor himself may share this sentiment; in his more recent work he has updated his view of how concepts are acquired, and has overtly rejected the need to posit innate concepts. In the next section I will critically assess this more recent proposal, and argue that the view is not as different from the current view as it is claimed to be, at least not in ways that matter to addressing the concerns raised above.

3.4 Fodor’s Mind-Dependence Proposal

Fodor (1998) revisits the question of concept acquisition, with an eye to offering an account of concept acquisition consistent with his view of concepts as unstructured atoms. While Fodor characterizes this account as an alternative to nativist views of acquisition, I will argue that his account shares a commitment to methodological nativism with his earlier view. The major difference between the new view and the older view is Fodor’s commitment to the mind-dependence of many lexical concepts.

Fodor is motivated to revise his previous account of concept acquisition because of the “doorknob/DOORKNOB problem”: that is, the problem of explaining why it is that concepts are so often acquired from experiences of things that fall within their extension. This, as noted in the previous section, is something of an embarrassment for triggering views of concept acquisition, since, as Fodor stressed, it is a core commitment of the triggering account that the concept-trigger links can be arbitrary. Because of this, the triggering account can’t explain why, in fact, these links are not arbitrary; why it is that concepts are in fact so often acquired from experiences of instances of them.
Confirmation-Theoretic models, as I argued in chapter 1, explain this by appeal to the fact that things within the extension of a concept are good sources of evidence about the satisfaction-conditions of that concept. The existence of this phenomenon, therefore, argues for learning models over triggering models, unless the latter have an alternative explanation available.

Fodor’s strategy is to treat concepts that are readily acquired from experiences with their instances as concepts that express mind-dependent properties. He proposes to treat concepts of artifacts, for example, as analogous to concepts of sensory properties: both are equally appearance concepts, concepts the contents of which are constituted simply by how things in the world strike our kinds of minds. The property that the concept DOORKNOB expresses is constituted by the fact that certain items in the world (those things that satisfy the doorknob stereotype; i.e. are typical doorknobs) cause us to acquire the concept DOORKNOB. There is no way to say what the property doorknob is – and so no way to say what the concept DOORKNOB means – over and above saying how typical instances of it strike us. So DOORKNOB works in the same way that RED does; the property that RED expresses is constituted by the fact that we take (typical) samples of redness to be instances of the same color. As Fodor puts it:

If you take seriously that DOORKNOB hasn’t got a conceptual analysis, and that doorknobs don’t have hidden essences, all that’s left to make something a doorknob … is how it strikes us…. My story says that what doorknobs have in common qua doorknobs is being the kind of thing that our kind of minds (do or would) lock to from experience with instances of the doorknob stereotype. (Cf. to be red just is to have that property that minds like ours (do or would) lock to in virtue of experiences of typical instances of redness)…. The basic idea is that what makes something a doorknob is just: being the kind of thing from experience with which our kind of mind readily acquires the concept DOORKNOB. And, conversely, what makes something the concept DOORKNOB is just: expressing the property that our kinds of minds lock to from experience with good examples of
instantiated doorknobhood…. [W]hat I want to say is that doorknobhood is the property that one gets locked to when experience with typical doorknobs causes the locking and does so in virtue of the properties they have qua typical doorknobs. We have the kinds of minds that often acquire the concept X from experiences whose intentional objects are properties belonging to the X-stereotype.” (1998: 136-8)

So, just as all that is required to have the concept RED, according to Fodor, is to have red things reliably strike us as red, so all that is required to have the concept DOORKNOB is to have doorknobs reliably strike us as doorknobs. The explanation of the d/D effect, in the case of concepts of mind-dependent properties, is thus a metaphysical explanation rather than a psychological explanation. Concepts are not reliably activated by their instances because of any feature of the mechanisms that mediate acquisition. Rather, they are activated by their instances because what it is to be an instance of such a concept simply is to have, as one of its characteristic effects on us, the result that we possess the concept in question.

So, in the case of mind-dependent concepts, the doorknob/DOORKNOB effect does not need a psychological explanation. Thus the failure of triggering views to provide an explanation of the doorknob/DOORKNOB effect does not count against them. So there is no additional problem generated by the doorknob/DOORKNOB effect about the acquisition of concepts that express mind-dependent properties.

On the other hand, noting this does not provide a solution to the problem of acquisition for such concepts. Consider, for example, the concept DISHWASHER. The property this concept expresses, according to the present view, is a mind-dependent property. That is to say that it is a property constituted by the effect that typical dishwashers have on human minds. We acquire DISHWASHER in virtue of interacting with
typical dishwashers. But then why is it that typical dishwashers cause minds like ours to acquire **DISHWASHER**? Even if the property of *being a dishwasher* were just the property *being the kind of thing that minds like ours would acquire from experiences of typical instances*, that still would not explain how it is that such experiences lead to our forming such concepts.

In fact, what Fodor claims is that the explanation of how such concepts are acquired may simply not be a matter for psychological explanation at all. “The moral of all this may be that though there has to be a story to tell about the structural requirements for acquiring **DOORKNOB**, intentional vocabulary isn’t required to tell it. In which case, it isn’t part of cognitive psychology.” (Fodor 1998: 143) While Fodor offers this conclusion as an alternative to nativism, the claim that concept acquisition is not within the explanatory domain of cognitive psychology fits the conception of methodological nativism outlined above.

If the thesis that many of our concepts express mind-dependent properties is supposed to illuminate the issue of how concepts are acquired, then it should be possible to explain how our minds come to be in such a position as to be affected by these properties in such a way as to enable us to acquire the concept. But, for Fodor, it is simply a brute fact that our minds are such as to be affected in such a way. But this kind of story is not made any more palatable when it is told about a mind-dependent property as when it’s told about some other kind of property.

---

38 For that matter, what makes something a typical dishwasher? There is a prospect of circularity here, since if what makes something a typical dishwasher is that it is typical for a dishwasher, and dishwasher is in turn constituted by reference to what typical instances of it would cause in us, then there is no independent explanation of the relevant notion of typicality.
In fact, the mind-dependence of properties would seem to be a red herring as far as a story about acquisition goes. The mind-dependence view makes a claim about the nature of the property that a given concept expresses. This may show that the link between concept and content is not arbitrary. But it does not explain in virtue of what it is that our minds grasp the concept in question. In fact, so far as the mind-dependence story goes, we might only be able to get our minds in a position to be properly affected by such properties by forming and testing semantic hypotheses about the application-conditions of the concept. The thesis that concepts express mind-dependent properties doesn’t decide among different theories of acquisition on offer.

Thus, whatever reason there might be for accepting the mind-dependence of the properties that many of our concepts express, such a thesis does not seem to illuminate the general question of how concepts are acquired. Thus Fodor’s proposal does not seem to represent an advance on the problem.

3.5 Conclusion

In this chapter I have argued that nativist accounts of the development of our conceptual repertoire are problematic. I argued that a distinctively nativist thesis of the origin of concepts needs to give a positive account of the nature of the processes that eventuate in concept possession. But the only developed proposal for this, the idea that concepts are triggered by experience rather than learned from experience, faces serious difficulties, difficulties which are not ameliorated by claiming that such concepts express mind-dependent properties. So nativist accounts of the origin of concepts seem unpromising. But we saw in chapter 1 that learning models of concept acquisition appear

39 As Fodor’s (2001) response to Cowie appears to concede.
unworkable as well. So the position we have reached is that there does not appear to be any tenable story about how concepts are acquired.

One response to this conclusion is a pessimism about explaining the origin of our concepts. Another alternative is to revisit the considerations from chapter 1 that led to this point; in particular, to try to work out an alternative to the Confirmation-Theoretic model of acquisition while retaining commitment to CMT, and thus to the psychological expicability of concept acquisition. The next chapter will address two strategies for pursuing the latter alternative. Ultimately, I will argue that neither of these strategies works. But I think their failures are revealing, and understanding where they go wrong will put us in a position to see how to explain concept acquisition as a psychological phenomenon.
CHAPTER 4  RECENT APPROACHES TO EXPLAINING CONCEPT ACQUISITION

4.1 Introduction

I have argued that the specter of radical concept nativism arises from a commitment to the Conceptual Mediation Thesis. Acceptance of CMT is typically manifested in a commitment to Confirmation-Theoretic models of concept acquisition, according to which the forming and testing of hypotheses about the satisfaction-conditions of concepts is the distinctively psychological mechanism of concept formation.

I argued in the previous chapter that the leading nativist candidate for explaining conceptual development faces problems of its own. This, combined with the intuitive implausibility that radical concept nativism is felt to have, has led to attempts to answer the arguments for radical concept nativism in ways consistent with acceptance of CMT. In this chapter I will examine two such attempts: Ned Block’s (1986) contention that a conceptual-role semantics for concepts yields an account of acquisition that avoids the problems raised for concept acquisition, and Eric Margolis’s (1998; Laurence and Margolis 2002) account of concept acquisition as the acquisition of a sustaining mechanism.

While these accounts differ significantly from each other, they have in common a rejection of the traditional conception of conceptual structure I set out in chapter 1, and of
the resultant notion of concept acquisition as a process of construction. Neither of these views require concepts to be semantically structured entities. But, although each view rejects the traditional expression of CMT, that they both remain committed to CMT itself.

I will argue that neither of these accounts succeeds in providing a generally applicable account of concept acquisition. While each has the potential to account for the acquisition of some restricted subset of our conceptual repertoire, neither has the resources to account for the acquisition of many of our concepts. And in each case, I argue, this is for substantially the same reason: a commitment to CMT. I take the difficulties these proposals face to provide additional evidence against CMT.

But my purpose here is not merely critical. Each of these proposals contain important insights which, when divorced from acceptance of CMT, point the way to a more adequate account of concept acquisition. So the results of this chapter will assemble materials toward a more adequate solution.

4.2 Conceptual Role Semantics and Concept Acquisition

Block argues that a conceptual-role semantics is well-placed to explain the acquisition of concepts and avoid the problems the arguments from chapter 1 raise for learning theories of acquisition. Understanding and evaluating this proposal requires first clarifying the commitments of conceptual-role semantics, as contrasted with more traditional semantic theories.
4.2.1 The Commitments of Conceptual-Role Semantics

“Conceptual-role semantics” refers less to a specific theory of the content of mental states than to a characteristic approach common to a family of distinct views. All these views have in common the claim that the identity of a concept is determined in some way by the relations it bears to other concepts.40 What makes a particular concept the concept that it is, is the way in which it is related to other concepts. Generally the relevant relations are those that characterize a particular functional role the concept plays in a cognitive system. This functional role, in turn, is a matter of inferential connections between the concept and others, as well as relations between the concept and perceptual input and behavioral output.

The identity of a concept, on conceptual role views, might be determined either by its total functional role in a cognitive system, or by some subset of its functional role. The thesis that the identity of a concept is determined by its total functional role is holism, while the thesis that the identity of a concept is determined by some subset of its total functional role is molecularism. (Fodor and LePore 1992; Devitt 1996)

Conceptual-role theories also differ as to whether they claim to give a complete account of mental content. One-factor theories regard conceptual role as a complete characterization of content. Two-factor theories hold that conceptual role characterizes one aspect of content, but does not characterize all there is to content. On two-factor views, conceptual-role content needs to be supplemented with a theory of referential content.

40 Conceptual-role semantics has its roots in the work of Wittgenstein (1953) and Sellars (1954; 1974). More recent versions of it are defended by Brandom (1994), Churchland (1990), Harman (1982), Loar (1981), and McGinn (1981). I will focus on Block’s version because he is the only one, to my knowledge, who has argued that conceptual role semantics addresses the arguments against concept acquisition.
Furthermore, conceptual role theories differ in whether they take the functional roles the theory specifies as constitutive of content to be *wide* or *narrow*. On the former view, the kind of content that a conceptual role semantics characterizes is wide content: content specified in terms of mind-external objects, properties and kinds. On the latter view, the kind of content that a conceptual roles semantics characterizes is narrow content: content specified in terms purely internal to the subject.\(^{41}\)

Block’s version is a *two-factor* version of *narrow holism\(^2\)* conceptual role semantics. On Block’s view, conceptual role characterizes the aspect of content that is internal to a cognitive system, while the aspect of content that is external to the system receives some kind of causal/informational account of content. But this is not to say that the two factors are completely independent. On Block’s view, the conceptual role factor determines which theory of reference is true of concepts: “what theory of reference is true is a fact about how referring terms function in our thought processes. This is an aspect of conceptual role. So it is the conceptual role of referring expressions that determines what theory of reference is true. Conclusion: the conceptual role factor determines the nature of the referential factor.” (Block 1986: 643-4) This is not to say that conceptual role determines reference, but it is to say that the way in which concepts refer to their objects is a matter of their having the functional role that they do.

---

\(^{41}\) These last two sets of distinctions are not unrelated, if one assumes that the purpose of ascribing content to cognizers is ultimately to explain how cognizers can represent mind-external objects, properties, and relations. One can either claim that this can be done directly, by specifying the conceptual role content in terms of those objects, properties, and relations (in which case one holds a wide one-factor version), or one can claim that the conceptual role content must be specified in terms internal to the subject, in which case one needs a further notion of content to explain how mind-independent entities can be represented (in which case one holds a narrow two-factor version.)

\(^{42}\) It is not clear whether Block means to endorse holism or molecularism here; on the one hand, he says that conceptual role is total causal role, but he also wants to make space for “collateral information”, information that isn’t constitutive of the identity of a concept.
What distinguishes conceptual-role semantics, in all its forms, from more traditional conceptions of the content of concepts, is that it rejects the distinction between *primitive* and *complex* concepts. On the picture sketched in chapter 1, complex concepts are those that are constituted by other concepts, while primitive concepts are those that are not. On the traditional view, there is some basic privileged set of concepts whose identity owes nothing to any other concepts. According to conceptual-role semantics, this is not true: all concepts are individuated, at least in part, through the relations they bear to other concepts.\footnote{Bill Lycan points out (in conversation) that it is open to the conceptual-role theorist to accept that there are some concepts that are complex in the traditional sense; viz., those that do have definitions of the traditional sort. This is quite true, so the claim above is better put as the claim that the kind of relations between concepts that CR recognizes are not, except accidentally, those that the traditional view recognizes. But nothing in the present discussion will turn on this.} Since the arguments from chapter 1 presupposed that concepts are acquired by using primitive concepts to generate complex concepts, a view of the content of concepts that rejects this distinction might have the resources to avoid the nativist conclusions of those arguments.

### 4.2.2 Conceptual Role Semantics and the Argument for Radical Concept Nativism

According to Block, conceptual role semantics offers the resources to resist the arguments for radical concept nativism. This is precisely because conceptual role semantics does not accept the primitive/complex concept distinction. This, according to Block, allows a proponent of conceptual role semantics to explain how concepts are acquired without appealing to complex concepts as traditionally conceived.

Block accepts a version of the Confirmation-Theoretic account of concept acquisition. That is, he holds that acquiring new concepts is a matter of forming and
confirming hypotheses about those concepts. Furthermore, he accepts that most lexical concepts do not have complex definitional structures. But, he argues, it would be a mistake to conclude from these two considerations that lexical concepts cannot be learned. That inference presupposes that learning new concepts must be a matter of hypothesizing the satisfaction-conditions for that concept. If the hypotheses by means of which a concept is acquired concern the satisfaction-conditions of the concept, then one acquires that concept by constructing it out of the concepts that one uses in the hypotheses that lead to its acquisition. And, if a concept is acquired by a process of construction, then the concept so acquired must be a structured representation. So, the thesis that concepts are learned by means of hypotheses concerning their satisfaction-conditions entails that learned concepts are structured representations.

The insight of conceptual-role semantics, according to Block, is that the hypotheses by which a concept is acquired do not have to be hypotheses about the satisfaction-conditions of the concept. The way we acquire concepts is not a matter of constructing them from other concepts; it is a matter of getting the concept to play a certain functional role. To acquire a concept is to acquire something that plays the functional role constitutive of that concept. But a concept does not need to be internally structured in order to play the functional role that it does. So getting something to stand in the right functional role does not require forming and testing hypotheses about its internal structure.

---

44 Though Block does not consider Fodor’s arguments against concepts having nondefinitional structure, nothing in the present dialectic turns on this. The kind of nondefinitional structure considered in chapter 1 still respects the distinction between primitive and complex concepts. So Block can accept that concepts do not have complex nondefinitional structure, in the traditional sense, either.

45 For the reasons I gave in chapter 1, I do not accept the reasoning embodied in the three previous sentences. I will return to this point in the next section.
How then do we get things to stand in the relevant functional roles? Block gives the example of learning basic physical theory to illustrate how this might work:

When I took my first physics course, I was confronted with quite a bit of new terminology all at once: ‘energy’, ‘momentum’, ‘acceleration’, ‘mass’, and the like. As should be no surprise to anyone who noted the failure of positivists to define theoretical terms in observation language, I never learned any definitions of these new terms in terms I already knew. Rather, what I learned was how to use the new terminology – I learned certain relations among the terms themselves (e.g., the relation between force and mass, neither of which can be defined in old terms), some relations between the new terms and old terms, and most importantly, how to generate the right numbers in answers to questions posed in the new terminology. (Block 1986: 647-8)

Learning new concepts is a matter of learning the relations among the new terms, as well as their connections to the concepts that you already possess. The relevant hypotheses involved are not hypotheses about the satisfaction-conditions of the target concept, but rather about the relation of that concept to other concepts, to perceptions, and to actions. This preserves the Confirmation-Theoretic conception of learning, and therefore preserves the idea that learning requires the mobilization of concepts one already has, while promising to avoid radical nativism.

4.2.3 Limitations of the Conceptual-Role Account of Acquisition

Block’s view, then, is that the view of acquisition made possible by adopting a conceptual-role semantics is importantly different from that made possible by the traditional view of concepts, and the difference allows this view to avoid commitment to radical concept nativism. Specifically, the conceptual-role view of acquisition does not
require hypotheses about the satisfaction-conditions about the concept being acquired, but instead appeals to hypotheses about the relation of that concept to other concepts.\textsuperscript{46}

Note that this view is committed to CMT, as his illustration of learning makes clear. Learning a theoretical vocabulary, such as the vocabulary of Newtonian physics, requires bringing some conceptual apparatus to the learning situation. To that extent, acquiring concepts depends on already having some concepts. Additionally, forming and testing the hypotheses to which Block appeals requires having concepts as well – those concepts that one needs to have in order to formulate the hypotheses in question. But this already suggests that Block’s view of acquisition is not different, in the relevant respects, from the picture that is threatened by arguments for radical concept nativism. If some concepts are required in order to tell the conceptual role learning story, then Block’s view is committed to a developmental analogue of the primitive/complex distinction: some concepts are required in order to explain how any others are.

This highlights a point I made in chapter 1: the thesis that concepts are structured is \textit{not} the pivotal claim in the arguments for radical concept nativism. On the traditional view, the process of acquiring concepts is conceived as a process of construction, and that thesis about acquisition is reflected in a claim about the end product of acquisition. The conception of acquisition as construction is reflected in the structure of the concepts thus acquired. But that the acquired concepts be structured is not a necessary commitment of the thesis that concepts are acquired Confirmation-Theoretically. The result of a Confirmation-Theoretic process might be the introduction of a new primitive symbol.

\textsuperscript{46} Block does not develop this proposal in any detail, so what follows will be based on attributions to the conceptual-role semanticist that goes beyond any explicit textual support. I think the proposals I consider are the most plausible way to extend Block’s proposal, but he should not be taken to be explicitly committed to what I go onto attribute to the conceptual-role view of acquisition.
This shows that the structure of the concepts thus acquired is not the relevant issue for whether the processes that lead to the acquisition of those concepts is committed to concepts one already possesses.

So the fact that the conceptual-role story of content is not committed to the traditional view of conceptual structure does not mean that it does not have analogous commitments to the traditional story of conceptual development. In virtue of its commitment to CMT, the conceptual role story of acquisition requires some developmentally primitive concepts to get its acquisition story off the ground.

This, of course, is not by itself an objection, nor does it imply a commitment to radical concept nativism. It simply raises the question of how many and what kinds of concepts the conceptual-role account requires to be developmentally primitive. Here the conceptual-role semanticist might have an advantage. The traditional view requires all internally unstructured concepts to be primitive. Since conceptual-role semantics does not require any concepts to be internally structured, it is not initially clear what concepts it requires to be developmentally primitive; this will depend on the details of the acquisition story it has to offer.

Notice, to begin with, that Block’s illustration is \textit{prima facie} not a promising model for concept acquisition in the general case. In general, those who are in a position to learn physics already have a wide array of concepts. Learning theoretical concepts, especially those referring to unobservable entities, is plausibly a later and more sophisticated kind of concept acquisition, one which requires a broad base of concepts already to be in place. For example, concepts of unobservable theoretical entities are not, in general, learned from experiences of their instances, precisely because they are
unobservable. It seems likely that, in order to acquire concepts of such entities, cognizers need to have concepts of the observable phenomena that, according to the theory, provide evidence for the existence and properties of those unobservable entities. Additionally, Block’s particular illustration has to do with explicit instruction in the conceptual framework of the theory. But the acquisition of a cognizer’s basic conceptual repertoire does not seem in general to depend on this kind of overt tutelege; we acquire many of our concepts without anyone’s instruction. While these points hardly invalidate the conceptual-role approach, they do raise questions about whether it is a promising model for the acquisition of nontheoretical concepts.

Nor is it entirely clear how to understand the role of the hypotheses Block identifies as involved in acquiring new concepts. Block talks about learning relations between concepts, both among multiple newly acquired concepts and between newly acquired concepts and already-possessed concepts. This suggests that the hypotheses one formulates in order to learn concepts concern these relations. For example, acquiring the concept FORCE (in Newtonian mechanics) might involve entertaining such hypotheses as “the concept FORCE is the concept of a value that is the product of the values of mass and acceleration.”

But this suggestion presupposes that the concept-learner already has the concepts MASS and ACCELERATION. But those are part of the cluster of theoretical concepts that are supposed to be acquired together. So possession of those concepts can’t be a precondition on acquiring FORCE. But how then should we understand the proposal?

One suggestion is that we can formulate such hypotheses in, so to speak, a purely formal manner. As the passage from Block above suggests, the claim is that we do not
have to learn new concepts purely in terms of those we already know. Rather, part of what we learn when we learn new concepts are their functional relations to each other. So perhaps the relevant hypothesis should be something like “the concept force is the concept of a value that is the product of the values of the concepts mass and acceleration.” The capacity to entertain this hypothesis does not presuppose possession of force, mass, or acceleration. What is entertained is a hypothesis about the relation between these concepts. Forming and confirming hypotheses about the total relevant relations between concepts is what counts as acquiring those concepts.

But this raises the question of how those hypotheses could be confirmed without employing the concepts of the theory. Recall that a central part of the Confirmation-Theoretic account has to do with the testing and confirmation of the hypotheses one formulates in the course of acquiring a target concept. This requires representing the evidence that bears on the hypotheses one forms. In order to do this, one needs to have the concepts that enable one to represent that evidence. This in turn means that the evidence required to confirm the hypotheses that result in new concepts being acquired must be fully representable in terms of concepts that one already possesses. Since the concepts being acquired are not internally structured, one is not confirming hypotheses about their satisfaction-conditions. Instead one is confirming hypotheses about their relations to other concepts. But it appears that what one confirms, in acquiring these new concepts, are that certain relations obtain between what are initially only formally characterized items.

Now suppose that it were possible to settle on the right network of relations among such items using only concepts that one already possessed. That would suggest
that the contents of the concepts thus acquired could be characterized in total by the concepts one already possessed. And that, in turn, amounts to a kind of operationalism about the concepts thus acquired.

As an example, consider again the concept \textit{FORCE}, and suppose that the concept-learner is in the situation that Block describes. Then the concept-learner will acquire \textit{FORCE} by forming and confirming hypotheses about the relation of \textit{FORCE} to other concepts. Now either these will be concepts that she already possesses, or they will be concepts that she is acquiring at the same time. In the latter case, her hypotheses will have to relate \textit{FORCE} to other concepts, such as \textit{MASS} and \textit{ACCELERATION}, that she also does not possess. So her evidence for hypothesizing one relation rather than another between those concepts will have to be representable in terms that do not use, but only mention, such concepts. Suppose, as is suggested by Block’s example, that this evidence consists of the correct answers to problems posed by the instructor. Then competence with these concepts amounts to being able to generate the correct output in response to give inputs. But, to claim that this is sufficient for possessing a concept is to suppose that all that is required for possessing a given concept is being able to generate the right outputs in response to the right inputs. But this is compatible with those concepts having no reference at all. So the account of acquisition provided by the conceptual role theorist is not an account of how concepts acquire their reference.

This is not, as it stands, an objection to a two-factor view of conceptual content such as Block’s, for it might be that the explanation of how our concepts acquire referential content is different from the explanation of how they acquire their conceptual-role content. On the other hand, however, this point shows that the account of acquisition
under consideration is incomplete, since it does not show how concepts acquire their referential content. So the conceptual-role account of acquisition needs to be supplemented with an explanation of how referential content is acquired.

Further reason to think that conceptual-role accounts of acquisition need to be supplemented with an account of the acquisition of referential content comes from a worry that conceptual-role semantics cannot account for the publicity of concepts. Fodor and LePore (1992) have claimed that holistic conceptual-role theories have the consequence that every change in the total mental state of a cognizer counts as a change in every individual mental state of that cognizer. On holistic conceptual-role views, the identity of a concept is constituted by the total causal/conceptual role of that concept. This includes the relations that it bears, directly or indirectly, to all other concepts the cognizer possesses. So, if any of those relations changes, the identity of every concept changes along with it. Now, acquiring a concept is, according to the conceptual-role account, in part a matter of altering the conceptual role of other concepts. So, acquiring a new concept implies that all the concepts the cognizer has post-acquisition are new. This implies that the concepts a child begins with, on the basis of which it acquires new concepts, are not concepts it shares with adult cognizers. Children’s concepts are incommensurable with adult concepts.47

Incommensurability need not, as such, be an objection to conceptual-role views.48 But it points out the need for some way of tracking what remains constant in such

47 Molecularist views of conceptual-role do not necessarily have this problem, since they hold only that some conceptual relations are constitutive of conceptual role. Such views, however, then face the challenge of explaining what relations those are and why they are so privileged.

48 Fodor’s worry about conceptual-role semantics does not concern development especially; the most pointed form of the worry is that, if any difference in conceptual role reflects a difference in content across the board, then no two people, and no two time-slices of the same person, will share any concepts, or
changes. A two-factor view has an answer here: if reference remains stable across changes in conceptual role, then conceptual roles can change over time consistent with there being a sense in which two people can have the same concept.

So there seems to be good reason to recognize referential content for concepts even for the conceptual-role theorist. If one wants to hold that concepts have two kinds of content, however, then one needs an explanation of the relation that these factors bear to each other. Fodor (1987) points out that it is problematic to hold that they are independent. If they are, then they might attribute different contents to a given concept, in ways that call into question the concept’s identity. Imagine, for example, that a particular concept stood in a referential relations to dogs, but played the conceptual role normally associated with a concept of chairs. Is this a DOG or a CHAIR concept? A two-factor theory seems to give no way to decide. On the other hand, if one of these kinds of content does have a privileged status in determining the identity of a concept, then what makes the other a semantically relevant feature of the concept? So a theory that recognizes multiple different kinds of content needs an explanation of how those kinds of content are related, so as to rule out these possibilities.

Block’s own view might seem to have a response to this problem. On that view, facts about the conceptual role of concepts determine what theory of reference is true. For example, the fact that referring terms function as rigid designators in our language (rather than abbreviated definite descriptions, for example) is a fact about how we use

Indeed any mental states at all. This result undercuts the generality of psychological explanation, since psychological explanations are general in virtue of applying to all people with the same mental states. I think, however, that the way in which two-factor semantics handles the developmental version of this problem also addresses this worry as well; if different cognizers can share reference, that is enough to save the generality of psychological explanation.
those terms; what role those terms play in our cognitive economy. More generally, the
fact that certain relations that obtain between mental representations and features of the
world count as relations of reference is settled by facts about the cognitive functions of
those terms.

But while this posits a relation between conceptual role and reference, it does not
seem to be an intimate enough relation to avoid the worry Fodor raises. This is because a
general relation between conceptual role and reference does not entail that any particular
conceptual role must be associated with any particular reference. The fact that we use
our referring expressions as rigid designators, and not as abbreviated definite
descriptions, is compatible with the situation described above, according to which a
concept has a reference severely at odds with its conceptual role. In order to rule out
such situations, there needs to be a reason to suppose that particular conceptual roles
cannot radically mismatch particular references, and the connection Block draws between
conceptual role and reference does not provide such a reason.\footnote{Bill Lycan suggests (in conversation) that there are possible cases of extreme mismatching between conceptual-role and referential content, such as the case where I am unwittingly transported to twin-Earth. I am not sure whether this would count as a case of mismatching according to the present view; since Block views conceptual role as narrow, the conceptual role associated with one’s WATER concept would be indifferent between, say, H$_2$O and XYZ. But however this may be, the potential mismatch is still restricted to natural kinds that have a different underlying microstructure but the same superficial features.}

I take these considerations to show, not that there is anything necessarily
intrinsically wrong with appeals to conceptual role, but that appeals to conceptual role are
incomplete as explanations of how concepts are acquired, because they do not address
how concepts acquire their referential content. So, even if concepts have conceptual-role
content, this fact does not by itself show how to avoid radical concept nativism.
I think this result is not surprising, since the conceptual-role account of acquisition is committed to CMT, and I have argued that commitment to CMT is at the root of the problematic features of theories of concept acquisition that generate nativist conclusions. The distinctive contribution of the conceptual-role semanticist was supposed to be that, because of his rejection of the traditional primitive/complex concept distinction, he was able to recognize different kinds of hypotheses that lead to the acquisition of concepts than those that involve the acquired concepts’ being constructed from the concepts used in the hypotheses. But, if I am right, then this is a relatively superficial difference from the traditional view of acquisition.

4.3 Sustaining Mechanisms and Concept Acquisition

Accounts of concept acquisition standardly assume that, if a concept is acquired by using other concepts, then the product of that process must be a structured concept. Whether structure is conceived on the traditional model as a matter of a concept’s being internally structured, or conceptual-role-style as a matter of the relations it bears to other concepts, claims that concepts are acquired by means of other concepts have been generally held to have implications for the structure of the concepts thus acquired. Given that assumption, Fodor’s arguments against most lexical concepts’ having structure are arguments against those concepts being acquired by a process of construction. It is this assumption that links claims about the products of acquisition with claims about the processes that lead to the acquisition of those products.

I gave some reasons for rejecting this assumption in chapter 1. If this assumption is rejected, then it may be possible to show how a process of using concepts to acquire
other concepts can result in the acquisition of new *primitive* concepts. If so, then rejecting this assumption may be the key to avoiding radical concept nativism.

This is the strategy pursued by Eric Margolis and Steven Laurence (Margolis 1998; Laurence and Margolis 2002). According to this strategy, acquiring new primitive concepts is a matter of constructing *sustaining mechanisms* for those concepts. Sustaining mechanisms are, as the name implies, mechanisms that sustain the relations that obtain between the concept and some aspect of the world in virtue of which the former represents the latter. To set up a sustaining mechanism is to put some internal item into a representational relation to some aspect of the world. Sustaining mechanisms are themselves structured entities, but they are distinct from the concepts acquired by means of constructing a sustaining mechanism. The Margolis-Laurence insight is that an unstructured concept can be acquired by a process of construction, if what is constructed is a sustaining mechanism for that concept, rather than the concept itself.

In this section I will critically examine the Margolis-Laurence view. I will argue that, while it offers several important insights, it ultimately fails to provide a generally adequate account of concept acquisition. I will argue further that this is because of their commitment to CMT.

---

50 For convenience, I will refer to this as the “Margolis-Laurence view”.

51 An early precursor of this view can be found in Fodor (1987). Cowie (1999) also develops some thoughts along these lines. Prinz (2002) offers a similar view, although it is distinctive in that the sustaining mechanism is the content-bearing item: the sustaining mechanism is the vehicle of the content (on Prinz’s view, this vehicle is a proxytype). So, according to Prinz, a concept qua proxytype can be structured even though it has a semantically unstructured content. Since nothing I say about the sustaining mechanism view will depend on the differences between these different views, I’ll phrase my discussion in terms of the Margolis-Laurence view.
4.3.1 Acquiring Concepts By Acquiring Sustaining Mechanisms

While Block argues for a view of acquisition on the basis of a distinctive view of the semantics of concepts, Margolis and Laurence intend their view not to depend on any particular theory of content; in particular, their view is intended to be compatible with a referential semantics for concepts that allows many concepts to be primitive. Their view is intended to show how, given a correct theory of the content of a particular concept, one can come to be in a position to have a concept with that content. Any theory of mental content, whether causal, informational, teleological, or otherwise, will specify some conditions that have to be satisfied in order for a concept to represent what it does. Given such a theory, the question about how concepts are acquired is: given the correct theory of mental content, how can we come to be in a state in which the conditions the theory specifies obtain?

For example, suppose that Fodor’s (1987; 1990) asymmetrical dependence theory of mental content is the right theory. This theory specifies conditions that a cognizer must be in to have a representation of a certain property. These conditions are rooted in the notion of a representation’s covarying with what it represents, as a matter of nomic necessity. To say that a representation covaries with what it represents is to say that the representation is tokened when and only when the feature of the environment that it represents is present. Covariance is not by itself sufficient for representation, since representations can be tokened correctly even when what they represent is not present. Fodor claims that such tokenings nevertheless represent that feature of the world because those tokenings would not occur unless tokens of the representation were produced when that feature is present. Roughly, the cognizer must be disposed to token the
representation in response to the aspect of the world that gives it its content, and any
dispositions that cognizer has to token the representation in response to other aspects of
the environment must be dependent on its disposition to respond to the content-conferring
aspect of the environment.

For example, to have the concept RABBIT, on Fodor’s view, is for tokens of one’s
RABBIT-concept to be disposed to covary with presentations of rabbits, as a matter of
natural law. But one’s RABBIT-concept isn’t tokened in response to all rabbits, nor are all
tokens of RABBIT produced in response to rabbits. What makes one’s concept a RABBIT-
concept (rather than a concept of RABBIT-OR-WHATEVER-ELSE-CAUSES-TOKENINGS-OF-
THIS-CONCEPT) is that the fact that other things can cause RABBIT-tokens depends on the
fact that rabbits cause RABBIT-tokens, but not the other way around. It is because rabbits
cause RABBIT-tokens that you can misidentify rabbits, or think about rabbits in their
absence, not because you can do these latter things that rabbits cause RABBIT-tokens.

Fodor’s view of content is atomistic in that for a representation to have the
content that it does is simply a matter of standing in the right nomic relations to the
feature of the world that it represents. Content is determined purely by non-dependent
nomic covariation. In particular, the content of a concept is not in any way constituted by
any other concepts, beliefs, or inferential dispositions the cognizer has.

However, the laws that are, according to the theory, supposed to be constitutive of
content require mechanisms to mediate the connection between the representation and
what is represented. It is not a brute fact that a lawful connection obtains between
representation and represented. It is not, for example, a brute inexplicable fact that
rabbits cause tokenings of RABBIT-representations. That such a relation does obtain is
due to other factors that mediate and sustain this nomic connection. It is, for example, because of complicated facts about the operations of our visual systems and how they deal with various stimulations that certain kinds of stimulations of those visual systems, in appropriate circumstances, lead us to produce a token of RABBIT.

Margolis and Laurence’s idea is to treat a cognizer’s beliefs and inferential dispositions as the mechanisms that mediate the nomic connections that are constitutive of content. Even though these aren’t constitutive of content themselves, they play a role in giving the concept the content that it has, by making it possible for the concept to stand in the relevant mind-world relationship that is constitutive of content. It is because we have certain beliefs about rabbits, for example that rabbits eat vegetables, that vegetables can sometimes lead us to token our RABBIT concepts. And it is because we are disposed to make certain inferences about rabbits, for example from the presence of a certain kind of track to the probable presence of rabbits, that the presence of certain kinds of tracks can sometimes lead us to token our RABBIT concepts. In these ways, other beliefs and cognitive structures that we have make it possible for some of these cognitive structures to stand in representational relations to features of the environment. These relations might be sustained in any number of different ways; while possessing a concept is simply a matter of having a representation under the control of a given property, that relation may be mediated by any number of different mechanisms. Different people may represent the same property in many different ways, compatible with their having the same concept. So different people may have different beliefs about rabbits, or be disposed to make different rabbit-regarding inferences, but so long as their rabbit-
regarding attitudes cause them to token their inner state appropriately, they will each count as having a RABBIT concept.

So possessing a concept is a matter of possessing an adequate sustaining mechanism for that concept. Similarly, acquiring a concept is a matter of acquiring an adequate sustaining mechanism. But, while acquiring a sustaining mechanism is a matter of acquiring beliefs and inferential dispositions, none of those play a role in constituting the content of the concept thus acquired. So the content of an acquired concept can be unstructured even though the processes used to acquire that concept involve other concepts, beliefs, and inferential dispositions. In order to have a RABBIT concept, you may need to have some beliefs of a certain kind, and be disposed to make some inferences of a certain kind; but the content of your concept may still be simply: rabbit.

Fodor’s argument turns on the apparent impossibility of learning a concept except by constructing it from other concepts. If this is right, then no primitive concept can be learned. But, if acquiring a concept is a matter of acquiring a sustaining mechanism to mediate the connection between a representation and its object, then a structured sustaining mechanism can grant the organism the capacity to acquire a new primitive representation. So, if learning a concept is a matter of acquiring a sustaining mechanism, then perhaps learning a concept does not require constructing it from other concepts.

So, in brief, Margolis and Laurence hold that concepts can be acquired by constructing sustaining mechanisms for those concepts. Contents are assigned to concepts in accordance with the dictates of the right theory of intentional content. The right theory of intentional content requires that an organism instantiate a certain pattern of dispositions, which pattern accords with the pattern of nomic dependencies the theory
says is constitutive of content. This requires the organism to have a sustaining mechanism to mediate the nomic connections. But sustaining mechanisms need not be unstructured; they can involve lots of associated beliefs and information that the organism possesses. So there’s no problem about acquiring a sustaining mechanism parallel to that of acquiring a concept by construction. So acquiring a sustaining mechanism can be a way of acquiring a genuinely new concept, even if that concept has an unstructured content.

4.3.2 Varieties of Sustaining Mechanisms

How, then, does a cognizer acquire a sustaining mechanism? According to Margolis and Laurence, this can be done in any number of ways. All that a sustaining mechanism has to do is make sure that the right connection between representation and intentional object is enforced. You might, for example, instantiate the right connection in virtue of knowing a theory about the object involved. Then your tokenings of the concept would be mediated by your knowledge of the theory, and you would be disposed to token the concept only when, according to the theory, a tokening is warranted. Suppose, for example, that you knew the relevant biological facts about what makes a rabbit a rabbit. Then you will know the conditions under which RABBIT is properly applied, and this knowledge will enable you to token RABBIT when rabbits are present, and not to apply RABBIT to things that are not rabbits.

This conception of the role of knowing a theory in possessing concepts incorporates an insight of the conceptual-role proposal discussed earlier. But it avoids the criticism I raised against the proposal there, because it offers an explanation of the
relevance of knowing a theory to the referential content of a concept. Learning a theory is a way of setting up the causal relations between representation and represented that underwrite relations of reference. So, as Margolis and Laurence note, appeals to conceptual role might be reconstructed in terms of sustaining mechanisms for conceptual content.

But knowing a theory is a special case; cognizers do not need to have a developed theory of the nature of a kind in order to have a concept of it. In the general case, one has a concept of something prior to learning a theory about it. This means that there must be other mechanisms capable of sustaining semantically relevant mind-world relations. The central one Margolis and Laurence identify is what they call a “syndrome-based sustaining mechanism”. The thought is that, in the general case, someone who has a concept of something will know some contingent but fairly salient information about the category to which the concept refers; that is, the individual will know what Margolis and Laurence call the relevant “kind syndrome”. Knowledge of this syndrome, in turn, will play the role of mediating the individual’s tokenings of her concept. For example, I may not know the underlying biological nature of rabbits, but I know various things about rabbits: that they are furry, have fuzzy tails and large ears, and eat vegetables. Because I know these things and others like them, I am disposed to token my RABBIT-concept when, for example, I see something furry with a fuzzy tail and large ears eating vegetables in the garden outside.

As it stands, this picture is incomplete, because my knowledge of the “rabbit-syndrome” encompasses only relatively contingent and superficial information about rabbits. If someone were to import rabbits from twin-Earth, nothing in my knowledge of
rabbits would enable me to distinguish them from earth rabbits. How is it, then, that knowledge of a syndrome allows me to have a RABBIT concept, rather than the concept RABBIT-OR-SOMETHING-ELSE-WITH-THE-SAME-SUPERFICIAL-FEATURES?

Margolis and Laurence claim that this gap can be filled by appealing to a child’s essentialist bias. (Gelman and Wellman 1991) There is evidence that children are predisposed to think that natural kinds have some essential properties in common, in virtue of which they are members of the same kind, and that, as a consequence, a nonmember of a natural kind cannot be made into a member by making superficial changes to it. If an individual is indeed equipped with such a bias, then, even though she may be disposed to token her RABBIT concept in response to things that are superficially similar to the kind but which do not share the underlying essence, she will also be disposed to withdraw her application of her concept to such things when informed that they do not share the underlying essence. Thus her application of her concept to instances of properties that her concept does not represent will be asymmetrically dependent on instances of the property it does represent.

Margolis and Laurence claim that this picture vindicates learning theories for concepts, because it appeals to a few general psychological principles to explain the acquisition of a large number of concepts. Given an essentialist bias and abilities to form various kinds of sustaining mechanisms, an organism can acquire semantically primitive concepts by generating sustaining mechanisms for them.
4.3.3 Problems With Sustaining Mechanisms

In order to evaluate Margolis and Laurence’s position, two claims should be distinguished. The first claim is that acquiring a concept is a matter of acquiring a sustaining mechanism. The second concerns which kinds of sustaining mechanisms are responsible for the bulk of concept acquisition in development. I believe the first claim is true and important, and reflects a welcome theoretical reorientation in approaching concept acquisition. So long as we think of the only possible way of acquiring concepts as through constructing them from other concepts, primitive concepts will be unacquirable. Thinking of concept acquisition in terms of the acquisition of sustaining mechanisms allows a way of thinking about the acquisition of primitive concepts without ruling out from the start that such semantically unstructured concepts could be acquired.

The second claim, however, is problematic. Margolis and Laurence sustaining mechanisms characterize sustaining mechanisms as collections of beliefs, knowledge, and inferential dispositions.52 For example, acquiring a concept by acquiring knowledge of a theory is a matter of acquiring a concept through the acquisition of beliefs and inferential relations. Similarly, acquiring a concept by means of acquiring a syndrome-based sustaining mechanism is a matter of acquiring a concept by acquiring beliefs about the typical, salient, and diagnostic features members of that kind tend to have.

52 Margolis and Laurence do not explicitly voice this commitment, but they characterize sustaining mechanisms as involving the acquisition of beliefs and information about the intentional object of the concept. For example: “The model [is] 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.” (Margolis 1998: 364); “A sustaining mechanism is a mechanism in virtue of which a concept stands in the mind-world relation that a causal theory of content … takes to be constitutive of content … The typical sustaining mechanism is cognitive or inferential.” (Laurence and Margolis 2002:37); “… the child collects and stores a range of information concerning broadly perceptual features of the object. If all goes right, this store comes to embody a kind syndrome: it incorporates information that is highly indicative of the kind and tends to be exhibited by the kinds paradigmatic instances.” (Laurence and Margolis 2002: 42) But beliefs, according to the Representational Theory of Mind, are themselves composed of concepts.
But acquiring beliefs and inferential dispositions requires one to have the concepts that make up those beliefs and that figure in those inferences. This means that Margolis and Laurence’s view is committed to CMT. They are concerned to reject the idea that concepts can only play a role in the acquisition of other concepts by constituting the concepts thus acquired. They also reject the idea that concept acquisition must be a matter of induction.53 But their view retains the core commitment to CMT, that concepts are acquired by means of other concepts that the cognizer already possesses. This commitment, I think, raises problems for their view.

Notice that Margolis and Laurence take the acquisition of natural kind concepts as their preferred illustration of how concepts in general can be acquired. This comes out in their favorable comparison of their own theory of concept acquisition with Fodor’s view of the mind-dependence of concepts discussed previous chapter. Margolis and Laurence claim that Fodor’s view amounts to a kind of defeatism; that it fails to offer any explanation of how concepts are acquired.

But, whatever the merits of this criticism of Fodor’s view, it does not favor Margolis and Laurence’s view over Fodor’s. That’s because this aspect of Fodor’s view is not a theory about how natural kind concepts are acquired. Fodor’s view is that natural kind concepts are a later conceptual sophistication, and acquiring them presupposes a large antecedent body of concepts. His theory about how natural kind concepts are acquired is explicitly very different from that of how other concepts are acquired. In fact, Fodor’s theory of how natural kind concepts are acquired is actually quite similar to

53 Although their reasons for this are not entirely clear. They dismiss this aspect of the argument for nativism very quickly, declaring that the issue of conceptual structure is at the heart of the argument. As I argued in chapter 1, I think that the issue of conceptual structure is in fact relatively peripheral, and that the conception of learning as induction is more centrally important.
Margolis and Laurence’s model. While Fodor claims that having a natural kind concept is just a matter of being appropriately locked to the property, he also claims that that locking probably has to be mediated by some sort of cognitive sustaining mechanism, such as knowing a theory of the property in question.\(^{54}\)

Noticing this highlights the extent to which Margolis and Laurence’s proposal is limited in its application to a restricted class of concepts. Consider again the kinds of sustaining mechanisms that Margolis and Laurence point to as candidates for acquiring concepts. All these mechanisms are explicitly cognitive, in the sense that they require other conceptual resources. To have a syndrome-based sustaining mechanism, for example, requires associating various pieces of contingent diagnostic information with the concept, and this presumably requires that the organism have the conceptual resources to represent this information. If, for example, one has the concept DOG in virtue of having a syndrome-based sustaining mechanism, then presumably the information associated with the concept includes something like the following: *is furry, has a tail, is four-legged, barks*, etc. But if the organism can use this information to control its tokenings of the concept DOG, then – given CMT – it does so in virtue of having the concepts IS FURRY, HAS A TAIL, IS FOUR-LEGGED, BARKS, etc. But those concepts are not themselves natural kind concepts, so Margolis and Laurence’s acquisition story does not apply to them as it stands. And on reflection it’s unclear how their story might be extended to apply to them. Take the concept HAS A TAIL. Suppose, along with Margolis and Laurence’s model, that to acquire this concept is to acquire a sustaining mechanism that mediates the nomic connection between the concept and the property it represents.

\(^{54}\) This line of Fodor’s thought is continuous with his (1987), where he discusses the role of knowledge of theories in mediating reference for concepts.
What kind of contingent and diagnostic information could mediate the relation between tails and TAIL? It is not clear what information could be used other than that which would be expressed using the concept TAIL itself; what is more salient and diagnostic of whether something instantiates the property having a tail than whether or not it has a tail? Having a tail is not a property that has a hidden essence; so there is no salient more superficial property that tells whether or not something instantiates the property than that very property.

The lesson here seems to be general. For a concept of any property that does not have a hidden essence, there is no distinction between contingent superficial properties and hidden essential properties. So there is no way to follow Margolis and Laurence’s model to generate a sustaining mechanism that does not presuppose possession of the concept that the mechanism was supposed to help us acquire. So Margolis and Laurence’s model does not extend to any concepts other than those that sustain an appearance/reality distinction. So their theory is restricted to explaining the acquisition of a relatively restricted range of concepts, and their story about how to acquire those presupposes that the cognizer already possesses a rich conceptual endowment.

I take these criticisms to show that, so long as sustaining mechanisms are taken to be conceptual mechanisms, appeal to the construction of sustaining mechanisms to explain the acquisition of concepts is of limited explanatory value. This is because appeals to the kind of sustaining mechanisms Margolis and Laurence describe require cognizers already to have a rich conceptual repertoire. So while the Margolis-Laurence view represents an advance on radical concept nativism, it still leaves unexplained how we acquire most of our concepts.
4.4 Conclusion

In this chapter I have examined two lines of response to the arguments for radical concept nativism. There is much in these views that, I think, is importantly right. Block’s emphasis on the inferential relations between concepts gives voice to an important intuition about the interconnectedness of conceptual representation, one that an atomistic referential semantics has difficulty capturing. However, that account is at least incomplete, since it does not provide an explanation of how concepts acquire their referential content.

Margolis and Laurence attempt to show how to integrate appeals to such inferential connections with a referential account of content, by thinking of such connections as mechanisms for sustaining content-constitutive relations. I think the notion of a sustaining mechanism is a welcome advance. But I think the characterization that they give of sustaining mechanisms limits their usefulness in explaining concept acquisition.

As I have argued before, I think the reason for the limitations of these views are their commitment to CMT. Commitment to CMT as a constraint on an adequate psychological theory of concept acquisition requires commitment to as many innate concepts as are needed to explain the acquisition of the rest. I have been arguing that, so long as we remain within this picture, we have the resources to explain at most the acquisition of a restricted subset of our conceptual repertoire.

On the other hand, as I argued in chapter 2, the use of representational structures is not the defining characteristic of psychological processes. This raises the possibility
that there could be nonrepresentational sustaining mechanisms for some of our concepts. If so, then acquiring those concepts might be a matter of acquiring such a nonrepresentational sustaining mechanism. On this picture, acquiring such concepts would not require the cognizer in question to have concepts already, and hence would not require her to have any innate concepts. The principles according to which our faculties of concept acquisition work may not depend essentially on the use of concepts we already have. In the next chapter, I will articulate just such an alternative.
5.1 Introduction

In the previous chapters I have explained how extant theories of concept acquisition are committed to the Conceptual Mediation Thesis, and how acceptance of that thesis renders explaining how concepts are acquired problematic. In this chapter I will offer an alternative account of concept acquisition that rejects CMT. The account I offer does not require the postulation of already-possessed concepts to explain how new concepts are acquired, and therefore requires no commitment to any innate concepts. My claim is that we do not need to postulate any innate concepts to explain concept acquisition.

Two questions must be addressed here. First, is there a coherent explanation of how a cognizer could acquire concepts without relying on representational capacities she already has? Second, is there reason to think that this is in fact how human cognizers do acquire concepts? The first is a question about the “design specifications” for cognitive systems: could we design a system that was able to acquire concepts in this way? The second is a question about psychology: are human cognitive systems such as to accord with this design? An affirmative answer to the first question alone would be enough to reject CMT. But I think the answer to both question is ‘yes’. I will develop an account of how cognizers could acquire concepts without having any concepts to start with, and I will suggest that this account is compatible with relevant work on the psychology of
concept acquisition. I cannot here provide anything like a full review of the relevant developmental literature, however, so my primary focus will be on giving an affirmative answer to the first question. Once we have a coherent account of how concept acquisition might proceed without appeal to concepts already possessed, we can then approach the experimental literature in a fresh light.

5.2 Representation and Indication

Before we can proceed, some clarification is needed. I have promised to offer an account of concept acquisition that rejects CMT, but I have not given a theory of what concepts are. To the extent that it is unclear what a concept is, it is unclear what CMT claims, and so unclear what rejecting CMT implies. It will be unclear as well what the end state of a process of concept acquisition involves, and so correspondingly unclear what conditions of adequacy an account of concept acquisition should meet.

Recall from chapter 1 that I promised to argue against CMT by arguing against the weaker Representational Necessitation Thesis:

(RNT) Necessarily, if a cognitive system acquires any concept, then it already possessed some mental representations prior to acquiring that concept.

Since RNT makes no claims about the role that these mental representations play in acquiring new concepts, nor does it require that any mediation be specifically conceptual mediation, taking RNT as the target rather than CMT avoids some of the
worries about unclarity I just mentioned. And since the falsity of RNT implies the falsity of CMT, arguing against RNT is a way of arguing against CMT.

But RNT itself stands in need of clarification. That is because RNT requires an account of mental representation more generally. I will not here try to canvas all the different ways in which the notion of mental representation might be spelled out. Instead, I will start with some plausible and generally accepted conditions on something’s being a mental representation, and draw out a notion of mental representation that meets those conditions.

Nearly everything responds to input from the environment in regular and predictable ways. Metal expands and contracts with changes in the ambient temperature, for example. Similarly, the states of organisms change in response to inputs from their environment. But that is not yet enough to say that such a state of an organism represents its environment. A human being sweats when the ambient temperature is high, and shivers when it is low; that is not to say that the sweating or shivering is a representation of the ambient temperature. At most it is a natural sign of the temperature. There is a difference between representing a state of the environment and merely reacting to a state of the environment. Reacting to external conditions is not the same as representing those conditions.

What further conditions need to be satisfied, beyond merely reacting to some environmental stimuli, in order for something to count as representing those environmental features? As I argued in chapter 1, mental representations are needed to perform two explanatory tasks: to explain the semantics of thought, and to explain the
influence of thought on behavior. Characterizing a system as using representations will therefore be appropriate when these explanatory tasks need to be discharged.

There are thus two distinguishable requirements on something’s being a mental representation: The first is that it have content; i.e., that it have conditions of semantic evaluation. On standard contemporary views of content, the fact that a mental representation has the content that it does is determined by the structure’s standing in the right relation to that aspect of the world that it represents. The nature of this relation is a matter of dispute: some argue that it is a causal relation, others that it is an informational relation, others that it is a teleological relation, and so on.\footnote{55 For references see chapter 1 note 3.} I will not enter here into the debates about the correct version of a theory of content, however, because the differences between such theories will not matter to the theory of acquisition I offer. The job of a theory of content is to explain the conditions under which some structure can be said to represent some feature of the environment. The job of a theory of acquisition, on the other hand, is to explain how a cognizer can bring a mental structure to stand in the right kind of relation. The theory of concept acquisition I will offer does not depend on the particular features of any candidate theory of content. It will depend only on features that are common to all such theories.

In order for something to have representational content, it must exhibit the phenomenon of intentional inexistence: that is, it must be possible for it to be directed toward something that does not exist. A consequence of this requirement is that it must be possible for anything that is to be a mental representation to be in error. Only
something for which there is the possibility of misrepresentation can count as representing at all.\textsuperscript{56}

The conditions in the previous paragraph are necessary conditions on something’s being a mental representation. But – despite the focus on misrepresentation in the literature on naturalized intentionality\textsuperscript{57} – they do not suffice for an internal state to be a mental representation. A sunflower can fail to track the position of the sun correctly due to clouds, mirrors, or artificial light, yet the sunflower does not track the sun by representing its position. Indeed, a cunning enough setup may make the position of the sunflower systematically miscorrespond to the position of the sun in the sky. But that it is possible for the sunflower to fail to correspond to the position of the sun does not mean that therefore the sunflower is representing the position of the sun (in this case, wrongly). In such a case, the mechanisms responsible for tracking the sun fail to do what they are supposed to. But that they fail in this way does not mean that they are succeeding in tracking something else, namely a representation of the sun, which representation happens to be incorrect. The failure in this case is on par with the failure of a toaster with a broken spring to pop the toast and switch off the heat at the appropriate time. This does not mean that the toaster misrepresents how much time is needed to appropriately toast bread, or misrepresents how toasty the bread is.

\textsuperscript{56} Perhaps these conditions apply in the first instance to \textit{propositional} representations, since those are the representations that make claims about the world, and are therefore evaluable as true or false. But concepts are the constituents of propositional representations, and the semantics of propositional representations derive from the semantic properties of their constituents. So, strictly speaking, what will be erroneous are particular tokenings of concepts in the context of particular propositional representations.

\textsuperscript{57} Most of the literature on naturalized intentionality has revolved around attempts to solve various versions of the “misrepresentation problem”: see Dretske (1986), Fodor (1987, 1990), Millikan (1989), and see also chapter 1 note 3.
Why should we think that the sunflower is not representing the position of the sun, despite the possibility of something like error in these cases? Because the position of the sun controls the position and orientation of the sunflower’s head. The sunflower has no capacity to track, orient toward, or anticipate the position of the sun when it is not present. This highlights an important feature of representations, that they are (or can be) used, and used correctly, in the absence of their intentional object. In the case of the sunflower, when the sunflower is not properly tracking the presence of the sun, it is then in error. But when genuine mental representations are produced in the absence of what they represent, they are not always done so in error. So the existence of “wild” tokenings that are not erroneous tokenings is a hallmark of representation.

This condition – what I shall call the stimulus-independence of mental representation – is necessary in order for mental representations to satisfy the second requirement on something’s being a mental representation that I identified above: that it exert the appropriate kinds of causal influence on the cognizer’s behavior. Specifically, mental representations explain the behavior of a system whose behavioral output cannot be explained simply by appeal to the environmental input of the system. For example, the angle of orientation of a sunflower varies with the position of the sun in the sky, but the sunflower’s behavior is (ceteris paribus) a strict function of the position of the sun. The flower cannot use any of its inner states to change its behavior. We can say, if we wish, that the position of the sunflower’s head represents the position of the sun, in that

---

58 Grush (1997) argues that the function of representations is not to provide an accurate reflection of the world, but rather to enable representation-users to perform operations like hypothetical reasoning. Grush thus makes a distinction between perceptual registration and mental representation; the former are concerned to give accurate information about the way the world is, while the latter are used to manipulate that information in order to determine how the world might be in response to various possible changes in circumstances. Similarly, Dennett (1974) argues that mental representations are valuable because they provide an “inner arena” to test potential actions and their consequences.
we can “read off” the position of the sun from the position of the sunflower’s head. But the sunflower does not itself use representations to guide its orientation with respect to the sun. In general, when a system is such as to have its behavioral responses controlled completely by the environmental inputs, that system is not a representing system. Representation-users can, as sunflowers cannot, engage in behavior directed toward an object that is not currently present to the system.\(^{59}\)

In a similar vein, Haugeland (1991; see also Clark 1997) suggests that we need to appeal to representations in order to explain the behavior of systems when those systems have to coordinate their behaviors with features of the environment that are not always reliably present to the system, when it does so by having something else “stand in” for that feature (in such a way that the stand-in is what guides behavior in the absence of the feature itself), and when that stand-in is part of a general and systematic scheme of representations. The intuitive motivation for these conditions is that, when we can explain the behavior of a system directly by reference to its environmental inputs, we do not need to invoke representations. Representational explanation becomes appropriate when the system’s behavioral output is not a strict function of its current environmental input, but rather when its behavior is guided by something that allows behavior to be directed at some aspect of the world that the system cannot currently interact with. Cognitive systems have goals, and in order to satisfy them, those systems must be able to form plans, engage in hypothetical reasoning, and adapt their behavior in appropriate ways toward the objects of their goals even when those objects are not currently present.

\(^{59}\) I have argued that it is a condition on something’s being a representation that the relevant system be able to use it to guide its behavior. Notice, however, that this is not the same as claiming that it is a condition on something’s being a representation \textit{with a particular content} that it be able to be used in certain ways. That claim is characteristic of teleological theories of content, but I remain neutral on that issue here.
In order to do this, they must be able to use their representations correctly in such situations. So it is a necessary condition on a system’s inner states being representations that the system can use them correctly in the absence of the objects of those states.

So mental representations, I claim, have two central features: (a) they exhibit intentional inexistence and can therefore be produced erroneously, and (b) they are stimulus-independent: they can be produced correctly otherwise than by what they represent. They need to have these features in order to satisfy the two explanatory goals for which mental representations are invoked: (i) to explain the semantics of thought, and (ii) to explain the causal role of thought in behavior. RNT, as I am understanding it here, claims then that it is a necessary condition on a cognizer’s acquiring anything that satisfies both (a) and (b) that it already have some things that satisfy both (a) and (b). My claim, in contrast, is that we can acquire at least some things that satisfy both (a) and (b) without having anything that satisfies them to begin with.

I am not claiming that there is no notion of representation that might fail to satisfy (a) and (b), nor am I interested in legislating the correct usage of the term “representation”. But I do claim that representations as I here characterize them play a central and important explanatory role in contemporary cognitive science, and that those representations to which theorists appeal to explain concept acquisition generally do satisfy this characterization.

There is one other caveat before I begin. I have promised to explain how concepts are acquired. But, as I argued in chapter 1, concepts are distinctive kinds of mental representations. That means that satisfying (a) and (b) are not sufficient for something’s being a specifically conceptual representation. So acquiring something that
satisfies (a) and (b) is not sufficient for acquiring a concept. However, since concepts are mental representations, satisfying (a) and (b) are necessary conditions for something’s being a concept. I will address the further issue of what is distinctive of conceptual representation in chapter 6. The results of that chapter will complete my explanation of where concepts come from. Here my goal is to explain how new mental representations more generally can be acquired.

5.3 Perception and Representation

I will begin to approach the question of whether concept acquisition requires representational mediation by considering a parallel claim about representational mediation in the theory of perception, paradigmatically in the theory of vision. There are several reasons for this. First, many of the issues that arise in connection with concept acquisition also arise in connection with perception. This is because perception comprises the mechanisms by which we acquire new information about the world, just as concept acquisition comprises the mechanisms through which we acquire new ways of representing the world. Second, perception is the channel through which we acquire new concepts. In order to understand concept acquisition, therefore, we need to understand perception. Third, discussing perception will enable us to isolate some resources that will become relevant to motivating an alternative conception of concept acquisition.

5.3.1 The Traditional View of Perception

The traditional view of perception conceives of it as an inferential process that operates on representations to produce new representations. On Marr’s (1982) theory of
vision, for example, perception proceeds via the construction of progressively more
detailed representations of the visual scene. The process starts with some basic
representations generated by basic processes of transduction, and proceeds to produce
more detailed representations (the raw primal sketch, the full primal sketch, the 2½D
sketch, etc.), through processes that combine these representations with representations of
assumptions concerning the relation of these representations to their likely source.

This view is motivated by the claim that any single perceptual experience is
ambiguous with respect to its distal cause; the proximal stimulus that constitutes the basic
data for a perceptual experience is consistent with arbitrarily many distinct configurations
of properties in the world. Thus the precise cause cannot be simply “read off” the
proximal projection. Instead, the basic data for vision must be processed and interpreted
in order to produce an unambiguous percept. As Shimon Ullman puts it:

The fundamental problem underlying the interpretation of structure from motion
is the ambiguity of the interpretation. That is, there is no unique structure and
motion consistent with a given two-dimensional (2-D) transformation … To cope
with this indeterminacy of structure, the interpretation scheme must incorporate
some internal set of constraints that rule out most of the possible 3-D
interpretations, and force a unique solution, which in most real cases is also the
veridical one. These constraints may also be viewed as a set of implicit
assumptions about the physical world which, when satisfied, imply the correct
solution … The visual process can be viewed as being composed of two opposite
mappings: visual encoding and visual decoding. In the first, the structure of the
visible environment is encoded in a visual message that reaches the eye. This
message is then decoded by the visual system to yield the original structure of the
environment. (Ullman 1979: 142-3)

This processing is conceived as a matter of inference, conceived in turn as a
species of computation. The idea is that the basic data for vision consist of some
privileged set of representational resources, which constitute the premises for perceptual inference. (Fodor 2004)

Since all inferences need premises, there must be some representations that are produced in perception that are not themselves produced via inference. These representations are produced noninferentially, via processes of *transduction*. (Fodor 1983; Shapiro 2004) Processes of transduction are those that convert the perceptual systems’ responses to external stimuli into usable information. This raises the question which representations can be produced via transduction. According to Fodor (1986; Fodor and Pylyshyn 1981), the representations can be produced directly are those that represent *nomic* properties. Nomic properties are those that can figure in statements of natural laws. So what is distinctive about the properties that can be directly transduced, without inference, is that they are properties that can figure in statements of natural law.60

The initial perceptual representations are produced via processes of transduction. But from where do the representations of the background assumptions come? According to some views (Gregory 1970; Churchland 1989; Noë 2004), this information is learned. But, on the other hand, it is difficult to see how such assumptions could in general be learned, since their role is to disambiguate an ambiguous initial stimulus. But if they were to be learned through perception, then the relevant disambiguating information must be acquired on the basis of other, equally ambiguous stimuli. But if those other percepts are themselves ambiguous, then it is difficult to see how they could play the role of

---

60 This must be meant merely as a necessary though not a sufficient condition. *Quark* is, presumably, a nomic property, but it cannot be directly transduced, or even perceived at all, by human beings. The claim must rather be that, if a property is not a nomic property, then it cannot be represented simply through transduction.
disambiguating other percepts. For this reason, it seems that the assumptions that are needed to play the role of premises in perceptual inferences must be innate. (Fodor 1983)

Similarly, opinions differ as to what kinds of information can be used to interpret or process the initial perceptual representations. On some views (Gregory 1970; Churchland 1980), nearly everything a cognizer knows can potentially affect the content of her perceptions. On other views (Fodor 1983), the background information that can inform perception is part of an isolated and dedicated system that is relatively insensitive to what else the cognizer knows. In either case, however, the account of perceptual processing is the same; the debate here is just over which background assumptions are available to interpret the initial data of perception.

Note the parallels between the traditional view of perception and the Confirmation-Theoretic view of concept acquisition. In each case, the perceiver/cognizer is taken to start with some basic representations, which are then worked up into some output representation via inferential processes. In each case, the perceiver/cognizer is taken to have the basic representational capacities innately, along with whatever further representational resources are required to interpret the initial stimuli. The major difference is that, in the perceptual case, the representations produced are token representations which the system is already assumed to have the capacity to produce, while in the case of concept acquisition, the representations produced are new types of representations. But the logic of the two models is the same.61

---

61 The logic of these models, in turn, recalls the logic of Chomsky’s “Poverty of the Stimulus” arguments. In all of these cases, the argument goes, the gap between the fragmentary initial data and the final competence must be filled by innate constraints.
5.3.2 Perception and Representation

If the traditional view of perception is correct, then it may seem that RNT must be correct as well. If perceptual representations are generated by the production of initial perceptual representations that are then interpreted by further representational resources, then anything with the capacity to perceive must have a number of representational resources already. And if concepts are first acquired via the operations of perception, then a cognizer could not acquire concepts without having the capacity to perceive. It would seem to follow, then, that no concepts could be acquired without the mediation of the representational resources deployed in perception.

However, the traditional view of perception argues for RNT only if the role that perception plays in the acquisition of concepts requires it to be representational; that is, only if that role essentially depends on the representational features of perception. But the mere fact that a process can be performed by the use of representations does not mean that it can only be produced by the use of representations, that it relies on representations essentially. To take a simple example, one could build a teakettle with a thermometer that measures the temperature of the water, and with a signaling device that reads the thermometer and, when the water reaches boiling temperature, activates the signal. Such a system would use a representation of the temperature to generate an appropriate output at the appropriate time. But teakettles can be, and more generally are, built to signal that the water is boiling, not by reference to a representation of the temperature, but by reliance on physical principles. In this case, the relevant physical principles are that water turns to steam as it boils, and that steam forced through a small opening with
enough force will cause a whistling sound.\textsuperscript{62} By building something that exploits these facts, one can build something that operates in just the way that the representation-using device does, but which does not itself use representations.\textsuperscript{63}

So there is, in general, a tradeoff between representational structures and design constraints; what can be done by a system via the use of representations can also be done by building the relevant capacities directly into the “hardware” of the system. Given this, how is it determined whether or not a system, especially a natural system, performs its tasks by using representations? As I argued in section 5.2, systems that use representations are capable of greater flexibility in their operations than those systems that do not. The teakettle that operates by measuring the temperature of the water and signaling when the temperature reaches some prespecified level can be fairly easily reset to signal when the temperature reaches some different level. The more common teakettle cannot be so easily recalibrated, however; in order to signal when the water reaches a different temperature, the whole system would have to be redesigned. This suggests that, where the system displays flexibility with respect to its inputs, it should be treated as

\textsuperscript{62} In a similar vein, van Gelder (1995) discusses two ways to solve the problem of ensuring a constant flow of steam from a boiler: (i) via a governor mechanism that measures the flow of steam and computes and executes corresponding changes to the regulating mechanisms; (ii) a governing mechanism that relies on the steam pressure to propel a mechanism that reduces the flow of steam as the steam pressure increases and increases the flow of steam as the pressure drops. The former performs the task by using representations, the latter does not.

\textsuperscript{63} My use of the term “exploit” here echoes Robert Wilson’s (2004) distinction between \textit{encoding} and \textit{exploitative} representation. Wilson’s view of encoding representation is roughly the traditional notion of representation that I described above; exploitative representation, in contrast, is a matter of a system’s relying on constancies in the environment rather than on internal encodings of them. In my own view, exploitative representation ought not to be considered a form of representation at all, but I cannot enter into this issue here. For my present purposes, it is enough to note that it is not representation in the sense of 5.2 above.
using representations in its operations. Conversely, that a system is inflexible with respect to its inputs suggests that it is performing tasks without using representations.⁶⁴

So that a task can be performed by using representations does not mean that it must be so performed. Similarly, the fact that a system performs a task by using representations does not mean that its use of representations is essential to the way it performs that task. In this section I will outline an alternative conception of perception to the traditional view. On the view I will offer, perceptual states mediate the acquisition of new concepts, but those states do not perform that role in virtue of being representational states. In doing so I take no stand about whether or how perception might be representational. My claim is that, even if perception is representational, those features of perception in virtue of which it is representational are inessential to the role perception plays in the acquisition of concepts. The way in which I will appeal to perceptual states to explain concept acquisition does not require that they be representational, and so does not require commitment to RNT. I argue that, whatever else perceptual states are, they are at least indicating states, and the fact that they are indicating states is all that is relevant to explaining the role of perception in the acquisition of concepts.

I will begin by explaining the relevant notion of a perceptual state, focusing on the human visual system. The visual system operates by entering into various states in response to various patterns of irradiation of the eyes’ photoreceptors caused by the visible features of the environment. These states of the visual system are systematically interrelated; there are the complex relations between the kind of state caused by a particular pattern of irradiation, and the related states that would be caused by certain

---

⁶⁴ Stabler (1983) defends a similar condition in considering whether or not grammars are mentally represented.
transformations of the pattern of irradiation. For example, the state of a visual system when presented with a red, 1” radius circle differs in systematic ways from the state that it will be in when presented with a green 1” radius circle, or again when presented with a red 2” radius circle. Changes in the environment correspond systematically to changes in the states of the perceptual system of the organism, in ways determined on the one hand by the effect the environment has on the perceptual systems, and on the other hand by the structure of those perceptual systems.

These states of perceptual systems, such as the human visual system, are *indicating states* in the sense of section 5.2: they are states that the system correctly enters only when caused to do so by a specific feature of the environment, with all other productions of them being incorrect. Furthermore, the indicating states of perceptual systems are systematically interrelated in ways that reflect the relations between features of the world. Perceptual indicating states are not simply like a row of bulbs in a console, that merely switch on and off when the relevant feature is present, but provide no more information about the feature other than whether it is present or not. Instead, the manner of indication itself provides indication about the character of the stimulus. The indicating states for two different shades of color, for example, do not simply indicate the presence of one or the other of those shades; they also reflect the relations – of brightness, hue, saturation, etc. – between the colors the presence of which they indicate. They do so in virtue of relations obtaining between the two states that parallel relations obtaining between the features of the world they indicate.

Figure 1 illustrates the relevant relations. An arrangement of objects in the environment (a) causes the perceptual system to go into a particular state (b), while a
different arrangement of different objects \((c)\) causes the system to go into a different state \((d)\). The state \((b)\) indicates the environmental arrangement \((a)\), and the state \((d)\) indicates the environmental arrangement \((c)\), as the horizontal arrows reflect. The arrangements \((a)\) and \((c)\) are alike and different in various ways; for example, while \((c)\) does not have the same arrangement as \((a)\), the arrangement in \((a)\) can be transformed into the arrangement in \((c)\) in a straightforward way. These relations of likeness and difference between these features of the environment is reflected in corresponding relations between the perceptual states \((b)\) and \((d)\). That perceptual states are systematically interconnected in this way is reflected by the vertical double arrow between \((b)\) and \((d)\).

I will say that the relations perceptual states bear to each other, in virtue of which they are systematically interconnected in the way characterized above, and in virtue of which they reflect relations that different features of the environment bear to each other, are relations of Similarity. “Similarity” is introduced here as a technical term; precisely what relations these are, and precisely how cognitive systems use these relations, are empirical questions for continuing research in cognitive science.

So, systematically related states of perceptual systems indicate systematically related features of the environment. But that is not to say that they represent those features of the environment; not all indication is representation. I will argue that

---

65 Although for convenience I have depicted these relations pictorially, the perceptual systems need not indicate features of the environment by having elements organized in the same way as those in the environment are organized. The correspondence may be more abstract.

66 It is an empirical question as well which perceptual states stand in Similarity relations to which other perceptual states. States that indicate shades of color will stand in Similarity relations to each other, reflecting how close or different the shades are in hue, brightness, etc. But what about states that indicate color, on the one hand, and those that indicate shape, on the other? Additionally, Similarity relations may cross the boundaries of perceptual systems; perhaps states that indicate shape in the visual modality stand in Similarity relations to those that indicate shape in the tactile modality, for example.
indicating states can lead to the acquisition of new representational states. But, again, I have not argued that perception is not representational. For present purposes, I am concerned simply to maintain that the features of perception I have isolated are the features that are relevant to explaining concept acquisition, and those features do not require that such states are representational states. Whether those states are representational or not, the fact that they are indicating states is sufficient to for them to play the relevant role in concept acquisition.

**Figure 1:** A representation of the relations between objects in the environment and the perceptual states that indicate them.
5.4 Concept Acquisition Without Representational Mediation

In this section I will argue that we can explain the acquisition of concepts by appeal to perceptual indicating states, whether or not those states are representational. Recall from section 2 the problem to be solved is that of explaining how it is that a cognitive system can come to have a mental structure that stands in an appropriate content-constitutive mind-world relation to a feature of the environment. That mental structure must satisfy the conditions on mental representations set out in section 1:

(1) that structure must have an intentional object, and it must be possible for the system to produce tokens of that structure incorrectly;
(2) the structure must be stimulus-independent: there must be tokenings of that structure that count as correct even when not directly caused by something the structure indicates.

Furthermore, this explanation must not appeal to any other states that satisfy those conditions. I will argue that appeal to systems of indicating states as described in the previous section provides the basis for just such an explanation.

5.4.1 The Components of Concept Acquisition

My explanation will appeal to three different kinds of cognitive capacities: (i) capacities to have indicating states, (ii) capacities to record indicating states, and (iii) capacities to group current and recorded indicating states. I assume that these capacities are underwritten by mechanisms in cognizers’ heads, but I will give here only an abstract
characterization of these capacities, since my purpose is to show what kinds of capacities a cognizer needs to have in order to acquire concepts. Further exploration of how these capacities are implemented in cognizers’ heads must wait for another occasion. Indicating states I have already discussed in the previous section; here I will briefly explain the latter two capacities.

I will suppose that, in addition to the capacity to have perceptual states of the kind I have described, any cognitive system able to acquire concepts will have some capacity to record its perceptual states. To say that a cognitive system can record its perceptual states is just to say that the system possesses some subsystem – a “recording device” – that, when the system enters into a perceptual state, is caused to enter into a state that indicates that perceptual state. Figure 2 illustrates this:

**FIGURE 2:** Indication relations (marked by arrows) between states of different systems: a state of the environment (a) is indicated by a pattern on the surface of the sensory receptors (b), which is in turn indicated by a state of the perceptual system (c), which is in turn indicated by a state of the recording device (d).
An arrangement of features of the environment (a) will be indicated by a pattern of stimuli on the sensory surfaces of a cognitive system (b). That pattern will in turn be indicated by a state of a perceptual system (c). Finally, that state of the perceptual system will be recorded by an indicating state of the recording system (d). (Note that the manner of indication in each case may be quite different: in (c), for example, the shape of the cubes might be indicated by the four equal lines that compose each element, while in (d) the arrangement of the elements might be indicated by the location of the upward slope along each line.)

Note that recording is not, as such, representing, and the records thus generated are not, as such, representations. The capacity to produce records of indicating states does not itself imply the capacity to make any use of them, any more than the capacity of a camera to produce images implies any capacity on its part to do anything with those images.

I will also suppose that the cognitive system will have the capacity to group together perceptual states, both current and recorded, on the basis of certain of their Similarity relations. As with recording, “grouping” is an automatic causal process; an appropriately functioning grouping mechanism will group states together automatically given the relevant input. I will say that a group of perceptual states are “appropriately Similar” to each other when they are grouped together by the grouping mechanism on the basis of the Similarity relations they bear to each other. (What relations the grouping mechanism in fact uses is, of course, an empirical question).

Note that grouping is not, as such, representing, nor does it require the use of representations in order to operate. This is because grouping is automatic, and, as I
argued in section 5.2, that a process is automatic means that it does not require representations in order to operate.

![Diagram](image)

**FIGURE 3**: relations among indicating states occasioned by visual perception of a squirrel: (a) the actual squirrel in the environment; (b) the impression that squirrel makes on the sensory surfaces of the system; (c) the perceptual state produced in response to the sensory stimulation; (d) the state of the system’s recording device produced in response to the perceptual state (arrows signify relations of indication).

5.4.2 Concept Acquisition in Action

I will now show how a cognitive system equipped with these capacities can acquire a concept. I begin with an example. Suppose a child not yet equipped with a SQUIRREL concept encounters a squirrel. That squirrel will cause a set of perceptual states in the child, as shown in Figure 3.\(^{67}\) The squirrel (a) produces a set of stimuli on the periphery of the child’s visual system (b), which in turn produces a perceptual state (c)

---

*\(^{67}\) Here and throughout, I will restrict my focus to visual states.*
that indicates the stimulation and its environmental source. The recording device in turn produces a state \( (d) \) that indicates that perceptual state.

(As Figure 3 suggests, that state of the recording device may not appear from the outside to bear much correspondence to that which it indicates. Whether there is a resemblance detectable to the outside observer is irrelevant to the question of how one state indicates the presence of another. What matters is how the system itself uses these states.)

![Diagram](image)

**FIGURE 4:** relations among indicating states occasioned by visual perception of a moving squirrel: (a) the actual squirrel in the environment; (b) the impression that squirrel makes on the sensory surfaces of the system; (c) the perceptual state produced in response to the sensory stimulation; (d) the state of the system’s recording device produced in response to the perceptual state (arrows signify relations of indication).

Next, suppose that the squirrel runs across the child’s field of vision. This single object – the squirrel – then presents a temporally extended and evolving appearance to the child. The location and orientation of the squirrel will change as it moves, as will the
shape it presents to the child. Thus the perceptual states the squirrel causes in the child will change from moment to moment, for example from that shown in Figure 3 to that shown in Figure 4.

The indicating states will all change systematically from moment to moment as the squirrel moves. In order for the child to keep track of the squirrel, she will have to be able to treat this evolving appearance as an appearance, or series of appearances, of the very same thing, despite the alterations in her perceptual states. She will have to be able to follow its path and reorient herself and her perceptual apparatus to the squirrel.

Explaining how the child can do this will require appeal to indicating perceptual states and the Similarity relations between them, along with mechanisms that enable the child to keep track of, orient toward, and attend to stimuli. But none of this requires that the child have any mental representations. Orienting towards and tracking stimuli are well within the capabilities of even relatively simple mechanical devices, such as those employing motion sensors.  

Next suppose that the child loses perceptual contact with the squirrel, and later comes into perceptual contact with another squirrel (or the same squirrel at a later time), as in Figure 5. Since squirrels tend to look relatively similar, the child’s perceptual systems will enter into states appropriately Similar to those entered into upon seeing the

---

68 In fact, a relatively insensitive perceptual system may be to the child’s advantage. If a perceptual system had to take account of every alteration over time in order to represent what remained the same through change, the computational requirements would quickly become overwhelming. It would be far more efficient simply not to be sensitive to every potential change in what is perceived. Strikingly, children seem to be insensitive to perceptual information in just these ways. Recent work in developmental psychology (Spelke 1990; Baillargeon 1993) has demonstrated that infants do not treat salient perceptual information such as color, shape, and size as relevant to the identity of an object, relying instead on its motion, coherence, and cohesiveness in space and time. Thus, the limitations of the perceptual systems may in fact aid the child’s ability to treat objects as the same over time, and are predicted by the nature of the task of acquiring concepts. I will return to this literature in more detail in section 5.7.
first squirrel, and which are recorded in the child’s recording device. The fact that these perceptual states are appropriately Similar will reflect the fact that they are caused by things that are similar, in ways that are (to the system) significant and interesting, in the world.

Recall that we have equipped the child with mechanisms for recording and grouping together perceptual states that the child has been in and is currently in, on the basis of their Similarity to each other. Once the child has actually entered into some perceptual states that it groups together as Similar, the mechanisms that subserve these
capacities will be able to group further perceptual states with these as Similar in the same way. Included in this grouping may be perceptual states that the child has never actually been in, so long as those states are sufficiently Similar to states that the child has been in. The perceptual states that the child does undergo provide a basis for its grouping mechanisms to project its groupings to new cases. So, multiple perceptual encounters with squirrels will provide these mechanisms with the initial perceptual states that they group together, which in turn allows them to establish grouping relations to further potential perceptual states produced by future perceptions of squirrels.

Having grouped a family of Similar perceptual states together, these inner mechanisms can then produce a new type of inner state, into which the child will enter whenever it enters into a perceptual state appropriately Similar to this group. Then these perceptual states will then mediate the relation between perceptions of a feature of the environment and productions of the same mental structure – this new inner state-type – in response. This new state – call it $\Omega$ – has a kind of autonomy from the perceptual states that generate it, in that a token of it can be produced without a token of any particular perceptual state being produced as well. The state $\Omega$ might, for example, be produced by a front view of a squirrel, or a squirrel in profile, or any number of different squirrel-produced perceptual states. (See Figure 6 below.)

There is nothing mysterious about introducing this new inner state-type, as an analogy will show. Imagine sorting a number of different objects by putting different-colored stickers on them. Since some objects may belong to more than one of the sorting categories, some objects may receive more than one sticker. Suppose now that a new color of sticker – say, orange – is introduced with the rule: whenever an object has one of
a specified number of particular configurations of stickers on it, it receives an orange sticker. The rule for introducing orange stickers is tied to stickers that an object has already received, and so, that an object has an orange sticker on it does not entail that the object has one particular configuration of stickers on it. The introduction of new inner states to be produced when any of a certain number of configurations of perceptual states are produced is no more mysterious than the introduction of a new color of sticker to be applied when any of a certain number of configurations of other stickers. And, for just the same reasons as before, the relevant rule need not be represented to be causally effective.

So we can suppose that a cognitive system can produce new types of states, such as \( \Omega \), that are tokened when one set or another of a group of perceptual states are produced. Tokens of these new state-types are not directly produced by perception, but are produced in response to what is directly produced by perception. Since these new states are not tied directly to current perception, the possibility arises that they might come to be caused by things other than what is currently perceived. For example, tokens of one such state-type might come to be caused by tokens of another such state, one that is brought on by other groupings of perceptual states. Tokens of \( \Omega \) might come to be produced in response to tokens of another such state, one that is produced in response to perceptions of acorns, for example. As the number of these new states that the cognitive system has the capacity to produce grows, so will the causal networks in which these states can stand.
FIGURE 6: The operations of recording devices can produce a new inner state type ($\Omega$), to which different states of the recording device can then give rise (a-c). Once this new inner state type has been produced, the possibility arises that it might develop other causes. For example, in (d), a recorded perceptual state of a tree has become causally linked with $\Omega$.

This possibility of tokens of these states coming to be produced by causes other than those indicating states on the basis of which they were initially introduced arises from a feature of recording devices – they can bring into existence things that can stand in new causal relationships, from inputs none of which themselves can stand in those relationships. This phenomenon is not unique to the kinds of recording devices I have discussed here, but is ubiquitous in nature. Consider a river that initially forms because of rainfall. The body of water arises purely from the drops of water that falls from the sky, and the natural processes such as erosion that result. These natural processes act as a
natural recording device, reshaping the environment into channels for the water to run. But once the environment is so shaped, the river itself has a kind of autonomy from the drops of water that initially led to its formation, and can perpetuate itself, and continue to exist even once the rain stops. Similarly, once a being introduces these new kinds of states, they can become self-sustaining and self-supporting, and continue to be called on and used correctly in the absence of perception.

For example, this state $\Omega$ might come to be caused by perceptions of the trees that one has seen squirrels frequent, or the nuts that one has seen squirrels carrying. $\Omega$ might also come to be caused by recorded perceptions of those trees, or perceptions of nuts that are relevantly similar to perceptions of nuts that accompanied perceptions of squirrels. That is, $\Omega$ might come to be caused by perceptual states that do not indicate the same thing as those that led to the creation of $\Omega$, or by states of the system’s recording device that are not perceptual states at all, or even by other states of the same new kind as $\Omega$, once the system has created others.

In that case $\Omega$ has become stimulus-independent, in the sense of section 1, for it can be brought about by events other than those perceptual events that initially led to its formation. It is no longer a state that merely indicates squirrels, but a state that represents squirrels. Furthermore, the system will then be in a position to combine this new state with other states of this new type that it generates, in order to form more complex representing states. Then the child has acquired a SQUIRREL concept.69

---

69 Slight complications arise from the fact that SQUIRREL is a natural kind concept; there is the possibility of things that have the same overt perceptual signs as squirrels but which have different underlying essential properties. One possible way of accounting for how a SQUIRREL concept represents squirrels, given such possibilities, is that children are predisposed to treat natural kinds as having an underlying essence: see Margolis (1998) for discussion.
Crucial to this proposal is that this new stimulus-independent inner state be able to be correctly produced without being produced by the presence of what it represents. This, in turn, requires a story about when a tokening of a state counts as correct and when incorrect. Answering this question is a task for a theory of mental content, and different theories of content will answer this question in different ways. Since I intend the story of concept acquisition I offer to be compatible with a variety of theories of content, I do not want to take a definitive position on the question of what makes some tokenings of a representation correct or incorrect. But any adequate theory of mental content must have some resources to explain how a tokening of a representation can count as correct despite not being caused by what it represents, since this manifestly is a feature of human representational systems. So long as a theory of content can offer some account of this distinction, my proposal is compatible with it.

Suppose, for example, that Fodor’s (1987, 1990) Asymmetric Dependence theory is the correct theory of content. On this view, an inner state $X$ will represent $x$s just in case $x$s cause tokens of $X$ and, for any $y$s ($\neq x$s) that cause tokens of $X$, the following counterfactual holds: were it not the case that $x$s cause tokens of $X$, it would not be the case that $y$s cause tokens of $X$ either. In this sense, the causal relation between $y$s and $X$ is dependent on the causal relation between $x$s and $X$ – and since the causal relation between $x$s and $X$ is not similarly dependent on that between $y$s and $X$, the dependence is asymmetric. This theory is intended to accommodate the intuition that the causal relation between a representation and its intentional object is in some way primary or privileged relative to the causal relations between the representation and other potential causes of it. So, on this view, a cognizer will have a representation of $x$ when she is disposed to
respond to $x$s by tokening $X$, and she will have dispositions to produce tokens of $X$ in response to $y$s only because she has dispositions to produce tokens of $X$ in response to $x$s. (Margolis 1998) So the distinction between correct and erroneous tokenings of inner states is to be understood in terms of the asymmetric dependence of the dispositions one has to produce tokens of $X$ in response to $y$s on the dispositions one has to produce tokens of $X$ in response to $x$s.  

If this is the correct theory of content, then the view of acquisition that I offer can explain how something comes to stand in the relevant pattern of disposition. On my view, the initial causes of (what come to be) inner representations are those things that the representation represents. Later on that inner representation comes to have more potential causes, but the fact that it does depends on the fact that it is caused by its intentional object, since it is only because it stands in causal relations to that intentional object that it exists at all. So the other causal relations that this inner representation stands in are parasitic on its’ standing in causal relations to its intentional object. So the view of acquisition I offer shows how to develop the pattern of dispositions that Asymmetric Dependence says is constitutive of content.

Suppose, alternatively, that the right theory of content is a teleological theory of content (e.g., Millikan 1984; 1989) On teleological theories of content, what makes

---

70 In fact this cannot be the right way to understand the distinction between correct and erroneous tokenings of a representation. It is a distinction between tokenings that are caused by the representation’s intentional object, and tokenings that are not. But, as I have argued, it is an essential feature of representation that there be correct tokenings of the representation that are not caused by what they represent. This distinction – between non-x-caused-x-tokenings that are errors and non-x-caused-x-tokenings that are correct – cannot be reconstructed by relations of Asymmetric Dependence. Fodor (1990a), in considering a similar issue, suggests appealing to teleological considerations; Fodor (1990b) suggests appealing to conceptual roles. The former suggestion turns the proposal into a version of a teleological approach to content, which I consider below; the latter suggestion will be addressed in the next chapter. For the moment I pretend that asymmetric dependence is correct, since the question at hand is whether my view of acquisition is compatible with it.
something a representation of \( x \)s is that tokenings of it are used by the relevant systems to track and gather information about \( x \)s. So what makes a tokening of \( x \) correct, rather than incorrect, is that it allows the mechanisms that use the representation to fulfill their functions in a non-accidental way.

If this is the correct theory of content, then the account of acquisition I offer can explain how something comes to be an inner state that the system uses in this way. When a cognitive system needs to gather information about some particular aspect of its environment, it will need to have some way or ways of identifying that aspect of the environment. The indicating states of perceptual systems to which I have appealed can play this role. So these states can be the initial ways of introducing an inner state that the cognizer uses to track or identify features of its environment. More sophisticated ways of identifying are also available, but in order for them to be useful they will have to be ways of identifying the same thing. This will require that whatever inner states the system uses to identify and track \( x \)s must be able to be caused by a wide variety of different environmental conditions and other inner states, consistent with their being correct tokenings - which is to say that these inner states will have to exhibit the stimulus-independence that I identified above as a condition on mental representations. Since I have shown how inner states can acquire this feature, I think my account of concept acquisition is consistent with a teleological theory of content as well.

So I have explained how a child might acquire a SQUIRREL concept without requiring any representational states prior to acquiring that concept. The perceptual

---

71 I am assuming that my account of acquisition is compatible with whatever appeal to teleology is needed to make out a teleological theory of content. While I have said little about teleology in setting out my account of acquisition above, I see no reason why my account should not be compatible with any of various appeals to teleology.
states to which I have appealed are merely indicating states, not representational states. And the mechanisms for manipulating those states – for recording them and for grouping them as Similar – are relatively simple and automatic mechanisms that operate only when activated by the relevant stimulation. No representational states need to be invoked to explain these operations. Instead, a system equipped with the capacities to have various systematically interrelated perceptual indicating states, and with mechanisms to record those states, can create new states that satisfy the conditions to be representing states, and that are appropriately related to features of the environment to count as representing those features.  

5.5 Explanatory Goals of Theories of Concept Acquisition

In chapter 1 I identified two explanatory goals that a theory of concept needs to address: the door knob/DOR RN problem and the qua-problem. It has been claimed to be a virtue of Confirmation-Theoretic accounts of concept acquisition that they offer solutions to these problems. Now that I have offered an alternative account of concept acquisition, I can explain how these problems can be solved without appeal to innate representations.

Note that I am not presupposing that the representational ability thus acquired is essentially perceptual in nature. It could be that the representation thus acquired is somehow tied to a particular perceptual modality. (Prinz 2002 holds that concepts are essentially perceptual in nature, being copies of perceptual states.) But it could also be that it is not, that the representation thus formed has no tie to any particular sensory modality. The claim that concept acquisition is perceptually mediated is compatible with each of various views of the relation between perception and conception.
5.5.2 The doorknob/DOORKNOB problem

Fodor (1998) argues that a theory of concept acquisition must explain the doorknob/DOORKNOB effect. That is, a theory of concept acquisition must explain why it is that concepts are so often acquired from experiences of members of the category to which the concept refers. Since it is a contingent truth that we do acquire concepts in this way, there must be an explanation for why it is that it is such a prevalent way of acquiring concepts. Why do perceptions of squirrels lead us to acquire SQUIRREL, rather than (for example) MOCKINGBIRD? On Confirmation-Theoretic views, the reason concepts are acquired from perceptual encounters with instances of them is that instances of concepts provide a good source of evidence about the satisfaction-conditions of the concept. But that point does not by itself require that the cognitive system apprehend their value as evidence; it just requires that instances in fact provide such evidence, and that the fact that they do play a causal role in the acquisition of the concept. On my view, the concepts that are acquired from particular perceptual encounters are responsive to the specific character of those encounters, in virtue of the perceptual states they cause. So SQUIRREL, and not MOCKINGBIRD, is acquired from perceptions of squirrels because squirrels, and not mockingbirds, cause perceptual states of the kind that indicate squirrels. But the Confirmation-Theoretic view requires the cognitive system to appreciate the evidential relations that obtain between instances of the category as evidential relations prior to acquiring the concept. In contrast, according to my view, perceptual encounters do provide evidence for the presence of squirrels, but it is not essential for the cognitive system itself to recognize this fact. That being so, explaining the doorknob/DOORKNOB effect does not require appeal to representations.
5.5.3 The Qua-problem

The *qua*-problem is a problem about why we form one concept rather than another from an encounter with something that is an instance of many concepts. Why do we form SQUIRREL rather than MAMMAL from perceptual encounters with squirrels, since every perceptual encounter with a squirrel is equally a perceptual encounter with a mammal? And does answering this question require us to attribute representations to the concept-acquirer in order to explain why she acquires one concept rather than others in such a situation?

I think that evidence from psychological research on the order of concept acquisition suggests an answer to this problem that is compatible with the view I offer here. Eleanor Rosch and her colleagues (Rosch et al. 1976) have demonstrated the existence of what they call a “basic level” of concepts, which appear to be acquired relatively early in development. Strikingly, these concepts appear to be those that represent categories that people tend to find most perceptually salient, in the following sense: people have far more information about basic-level categories than they do about more abstract categories in the same hierarchy, but do not have significantly more information about less abstract categories in the same hierarchy than they do about the basic-level category. For example, people by and large know much more about squirrels (the basic-level category) than they do about mammals (the more abstract category), but not much more about eastern gray squirrels (the less abstract category) than they do about squirrels. Furthermore, the information that distinguishes the basic level from more and less abstract levels is by and large perceptual information: “squirrel” is a basic-level
category in large part because people know much better what squirrels look like than what mammals in general look like, but do not know much more about what eastern gray squirrels look like than they do about what squirrels more generally look like. Mammals as such have fewer and more abstract perceptual properties in common. This suggests that concepts of categories that are especially perceptually salient are acquired prior to concepts of more and less abstract categories. There is no necessity to this; presumably we are endogenously biased to acquire concepts of particularly perceptually salient categories relatively early in development. This bias provides an answer to the qua-problem, by providing a principled reason why concepts of some categories rather than others are acquired from experiences of exemplars of many different categories, and it provides just the answer we should expect if my view is right, since my view predicts that perceptually salient categories are those of which we are likely to form concepts first.\footnote{I have said little about acquiring concepts of superordinate categories here. The fact that such concepts tend to be acquired relatively later in development than “basic-level” concepts suggests that they may be poor candidates for being acquired directly from perception; perhaps acquiring concepts such as MAMMAL or FURNITURE requires that you already have some more specific concepts of things that fall under these superordinate categories (such as SQUIRREL or TABLE). I can concede this, however, so long as the concepts that one must have in order to acquire concepts of superordinate categories are concepts that can be acquired in the way sketched above.}

5.6 Empirical Support

In this section I discuss several lines of evidence from developmental psychology that bear on the question of how concepts are acquired. Some of this evidence has been offered as evidence that human cognizers do in fact rely innate concepts to acquire other concepts. Nevertheless, I will argue, these interpretations are undermotivated, and the available evidence is at least consistent with, and in some cases positively supports, the claim that we need not rely on concepts in order to acquire new concepts.
5.6.1 *The “Object Concept”*

There is mounting evidence that even very young infants expect the world to operate in regular and predictable ways. This evidence has been interpreted as supporting the idea that infants possess an innate object concept.

The seminal experiments are due to Renee Baillargeon (1993). Baillargeon habituated infants to a display with a screen rotating backwards and forwards through a 180-degree arc. Once habituated to this display, infants were exposed to one of two further events: a “possible” event or an “impossible” event. In each, a box was placed behind the rotating screen. The box was small enough to be completely occluded by the screen once the screen was completely upright. In the “possible” event, the screen continued to rotate backwards and stopped at the point where it would have impacted the (now occluded) box. In the “impossible” event, the screen continued to rotate on the same path as in the initial habituation event.

Children exposed to the impossible event looked longer at it than children exposed to the possible event. This suggests that the children were surprised by what had happened, which suggests in turn that they expected to see the motion of the screen change in consequence of the box being placed behind it. That, in turn, suggests that the child inferred that the box would affect the motion of the screen, and therefore that the child represented the existence of the box and likely effect on the screen. This would mean that the child has the capacity to represent physical objects.

Baillargeon argues that the impossible event is perceptually more similar to the habituation event than the possible event is, because the motion of the screen is the same
in both cases. So, if infant expectations were guided purely by perception, she argues, infants should be more surprised to see the possible event. The fact that they are not, then, means that they are acting on the basis of their representations of physical objects rather than simply on their perceptions.

Baillergeon and others have concluded from these and similar results that children represent the world in terms of physical objects from a very young age, and therefore likely have the concept innately. According to this view, children conceive of the world in terms of physical objects obeying the principles of cohesion, contact, and continuity.

I suggest instead that this evidence gives us reason to revise our conception of infant perceptions, rather than to attribute an innate object concept to the child. We should indeed reject the account of perception that Baillergeon criticizes. But the alternative to this need not be to attribute innate representations to the child. Instead, a more sophisticated account of perception can accommodate the data.

We can accept that infants are equipped to perceive the world in terms of physical objects obeying the principles of cohesion, continuity and contact. But we need not take this to show that infants represent these principles. They might instead exploit these principles in the course of perception. Their expectations might be a matter of being built to perceive in accordance with certain constancies, rather than representing those constancies to interpret their sensations.

Why might we prefer the second account to the first? The experimental evidence shows that children are surprised by violations of expectations in perception. Infants are surprised at impossible events when they are presented with them. But this does not give evidence that children have the capacity to represent physical objects in the absence of
perceiving them. It does not show, that is, that children are equipped with the capacity to make use of stand-ins for physical objects outside of perception. This suggests that children do not need the capacity to represent physical objects in order to perceive the world in those terms. In other words, the expectations children have about the behavior of physical objects do not exhibit the stimulus-independence characteristic of representation.

A crude notion of perceptual similarity is inadequate on independent grounds. It is true that the motion of the screen is more similar in the habituation and impossible event conditions. But to say that therefore the two scenes are themselves more similar, in the sense relevant to perception, is to take too narrow a view of what similarities the perceptual systems might be built to appreciate. The relations of similarity our perceptual systems are built to appreciate is not necessarily manifest from introspection.

Children may perceive the world in terms of physical objects rather than sense-data. That does not mean that they represent the world in terms of physical objects, or have concepts of physical object; nor are we forced to admit that they do so by the experimental evidence arising from this research program. A richer conception of perception allows an explanation of this evidence compatible with the denial of innate concepts of physical objects and principles to children.

Further support for this picture comes from Milner and Goodale’s (1996) hypothesis that there are two distinct streams of visual processing in the brain, the “what” and “where” systems. The former is a relatively slow processing system devoted to identifying the properties that objects have, while the latter is a relatively fast system devoted to locating objects in space. The former subserves visual recognition, while the
latter subserves tracking. These two aspects of vision are dissociable, which suggests that tracking objects does not require representing them or their properties as a condition of tracking them. If this is right, then it provides further support for the idea that perceptual systems might be able to track objects in perception without relying on representations.

5.6.2 Psychological Essentialism

Further evidence relevant to evaluating the view of concept acquisition I have offered here comes from studies of whether or not children are essentialists – that is, whether they believe that things are the same that appear similar on the surface, or whether they instead categorize things as similar or different based on deeper, more fundamental commonalities. Adult concepts seem to have some essentialist features: adults by and large believe that whales are mammals, not fish, for example, even though whales share many of the superficial features of fish. This belief is based on the idea that certain kinds of physiological and etiological characteristics matter more for a whale’s being the kind of thing that it is than do its external features.

Keil (1989) found evidence that second-grade children appreciate this distinction, and Gelman and Wellman (1991) report evidence that children as young as 4 years may appreciate the significance of the inside-outside distinction for something’s being a member of a certain category. For example, they found evidence that 4-year-olds think that changing the insides of an animal would tend to make it a different kind of animal, while changing the outside of a thing would not make it a different kind of animal. These
results suggest that 4-year-olds think that the insides of something have a privileged status when it comes to determining what kind of thing a thing is.

If this is right, then it may seem to put pressure on the idea that perceptually-based sustaining mechanisms can account for children’s concepts. If children are psychological essentialists, then the content of their concepts are not exhausted by what is available in perception.

In response, remember that my claim is not that concepts are constructions out of perceptual primitives. So my claim is not that the content of concepts reduces to the content of sensory concepts. My claim is that initial concept acquisition proceeds by mobilizing sensory states as sustaining mechanisms for the relations that constitute representational content. So it is compatible with this idea that what is represented – what the child’s concept is a concept of – is something that is neither a perceptual property nor a collection of perceptual properties.

What would be a problem, from my perspective, would be if conceptual resources were required to explain how children could be psychological essentialists. That is, if psychological essentialism is true, and if, in order for it to be true, children have to have representational resources prior to acquiring concepts of the kinds about which children appear to be essentialists, then my view is wrong.

However, I do not think that admitting that children are psychological essentialists does require positing representations to explain the acquisition of those concepts. Recall the discussion of the principles governing the OBJECT concept in the previous section. There I argued that, for a system to act according to such principles does not require that it represent them. Instead, a system might simply be built to act in accordance with those
principles, without using representations of them. And the absence of behavioral flexibility gives us reason to suppose that the system actually is built like that, rather than built to use representations.

In the present case, the principles in question are those that cause a child to treat something as dependent for its identity on its insides or not. Since children do not treat the identity of everything they encounter as dependent on its insides for its identity, they must have some way of identifying the things that do deserve such treatment.

The proposals in the literature for how children do this appeal to various kinds of predispositions or biases that children have, such as the shape bias, which causes children to treat certain things as having essences depending on their shapes. (Markman 1989) Similarly, children may be biased to treat things in the environment as living if they move without being moved by something else. These biases are posited to explain how children go from the perceptually available properties of a thing, such as its shape or motion, to some judgment about its important or essential properties.

This way of putting the point suggests that the child infers that a thing is of a certain kind on the basis of her current perceptual evidence together with a background assumption about the relation of that kind of evidence to how things are. Talk of “assumptions” the child makes, or “expectations” the child has, may encourage such a picture, since this language is of a piece with the language of propositional attitude ascription, and propositional attitude ascription is often taken to involve commitment to inner representations. But in fact these cases can be treated in much the same way as that proposed to treat perception more generally earlier in this chapter. These biases or predispositions need not be treated as representational structures that figure in inferences,
and, as before, the relative rigidity of these biases suggests that they should not be treated representationally, for the reasons I have given above.

So that acknowledging that children may be biased to think that certain imperceptible features of things are centrally important to what they are is compatible with the view of concept acquisition I have offered. Even when the content of a concept is not intrinsically perceptual, the mechanisms that mediate the acquisition of that concept might still be perceptual.

5.6.3 Sensory Vocabulary and Sensory Information

The evidence I have examined in the previous sections may seem prima facie at odds with the view of acquisition I have offered, and correspondingly it may seem as if the best I have done is shown that this evidence is not incompatible with the view I have offered. Even if this is true, it is not necessarily damaging to my position here. In the first place, I have been primarily concerned to show the coherence of the idea that one could acquire concepts while having none to begin with, and thus to show that it is a live empirical issue whether there are any innate concepts or mental representations. This much already provides us with a fresh lens through which to view the available developmental evidence. In the second place, I think there are general theoretical virtues to recommend the view I offer, including parsimony in how sophisticated a picture it requires of the innate initial endowment of the child, and explanatory generality in its promise to give a complete and unified account of the acquisition of mental representations. However, I do think there is some available evidence which is problematic for standard accounts of concept acquisition, and which my view better
accommodates. The available evidence of childhood lexical development suggests that children do not acquire their concepts on the basis of other concepts.

If concept acquisition is to be mediated by conceptual or representational resources that one already possesses, then it is plausible that among those representations that one must have to acquire any others are representations of sensory properties. These are a natural fit for a position that holds that there must be some unacquired representations, especially if those representations are to mediate the acquisition of concepts through perception. So acceptance of RNT naturally suggests the view that representations of sensory properties are among the unacquired concepts that underwrite the acquisition of all others. And in fact, most views of concept acquisition have held that sensory representations are not acquired.

One might also hold it to be a plausible principle that the order in which concepts are acquired is nonaccidentally related to the order in which words are acquired; specifically, that words that are acquired relatively later in linguistic development correspond to concepts that are acquired relatively later in conceptual development. However, this principle turns out to be at odds with the idea that sensory concepts are acquired early in conceptual development. Children begin to vocalize around the beginning of their second year, and at around 18 months their vocabulary begins to expand exponentially. The words children acquire in their second year are predominantly nouns, especially nouns for “basic categories”. Verbs are acquired later, and adjectives are acquired later still. Significantly, words for sensory qualities are among the latest words that children master; children do not become competent with words for colors until around 4 years of age, for example. (Carey 1978)
These facts seem puzzling on traditional views of concepts. On many standard views of language learning, learning a word is a matter of associating it with the appropriate mental representation. But if sensory concepts are part of the initial representational base, then it is puzzling why children find them so difficult to learn.

I am not claiming here that possessing a concept or a representation is a matter of mastering the use of a word, nor am I claiming that it is a necessary condition on having a concept or representation that one be able to use the corresponding word. However, if one has the relevant sensory representations to begin with, and especially if they are available to mediate the acquisition of other new concepts, then there seems to be no reason why children would not easily acquire sensory vocabulary. Theories of concepts that accept RNT and hold that sensory representations are part of the initial base should predict that these words could be easily learned. After all, it does not seem to be the case that children lack for evidence or stimulation in the world that would be relevant to their acquiring these representations.

Of course there might be other explanations for why this vocabulary is learned later than other vocabulary. But children’s limitations with sensory information are not limited to the late acquisition of words for sensory qualities. Young children seem to have limited capacities to make use of sensory information. Xu and Carey (1996; 2004) argue that infants do not make use of featural information as a guide to object identity. Infants are surprised when two objects are placed behind a screen, which is then raised to reveal only one object. But children are not surprised when one object is removed from behind a screen and replaced, and then a different-looking object is removed and replaced, when the screen is then raised to reveal only one object. This suggests that
children’s cues to the identity and persistence of objects do not depend on representing the perceptible features that the objects have. The ability to use such information to discriminate objects is a later sophistication.

If children are limited in the extent to which they can use sensory information, that suggests that they do not represent that sensory information as a precondition of representing objects. This does not mean, of course, that children are sensorially deficient in any way; children can perceive the colors of objects prior to learning the names for those colors. But it does suggest that children do not, and do not need to, represent the sensory qualities of objects in order to represent the objects. We represent objects first; our perceptual systems allow us to represent objects by reacting to the sensory qualities that objects have, not by representing those qualities themselves.

5.7 Conclusion

In this chapter I have presented a view of how at least some of our concepts can be acquired that does not require us to have any mental representations to begin with. I have therefore shown that RNT is false, and therefore that CMT is false as well. These results point the way to an account of where concepts comes from that does not require appeal to innate concepts, and thus points the way to a fully general account of concept acquisition.

The next chapter will address two remaining issues. First, as I pointed out at the beginning of this chapter, I have not here distinguished here between concepts and mental representations more generally. This opens the possibility that my account has not captured what is distinctive of specifically conceptual representation. I will address the
distinction between conceptual and nonconceptual representation, and show how the distinctive features of conceptual representation may be accommodated within the framework I have offered.

Second, I have argued in this chapter that concepts (or at least representational structures) can be acquired without the mediation of other representational structures. But I am not arguing for this claim in its full generality. That is, I do not argue that just any concept could be acquired by any cognizer, regardless of what other concepts that cognizer had, or whether it had any other concepts at all. This is unmotivated as well as psychologically implausible. Nor do I deny that concepts we already possess can be, and often are, involved in the acquisition of new concepts. On the other hand, traditional conceptions of conceptual mediation as a process of construction are, I have argued, untenable as well. So we need a new conception of how concepts mediate the acquisition of other concepts, when they do. I will offer an explanation for how this process works.
CHAPTER 6  CONCEPTUAL REPRESENTATION AND CONCEPTUAL MEDIATION

6.1 Introduction

In the previous chapter I offered a theory of concept acquisition according to which cognitive systems can acquire new representational abilities via the use of perceptual indicating states, recording mechanisms, and grouping mechanisms. This account provides an explanation of how cognitive systems acquire concepts that does not require the cognitive systems to have any innate representations at all. Therefore, it shows how to avoid the dialectic leading to radical nativism that I canvassed in chapter 1.

In this chapter I address two remaining issues. In section 6.2, I explain the distinction between conceptual and nonconceptual mental representations. I drew attention to this distinction in chapter 1, where I argued that, although concepts are mental representations, they are mental representations of a distinctive kind. I noted certain features that seem to distinguish concepts from other mental representations. In this section I will explain the underlying basis for those distinctive features of conceptual representation. In doing so, I will explain how inner states could come to have those features. This will complete the explanatory task begun in the previous chapter. There I showed how to acquire new mental representations; here I will show how those mental representations may become conceptual. I argue here that the capacity to have conceptual representations is intimately related to the capacity for metarepresentation: the ability to represent one’s own representations is what makes one’s own representations conceptual.
In section 6.3 I return to the question of the role of representational / conceptual mediation in concept acquisition. I have argued to this point that not all concept acquisition is mediated by representational or conceptual resources we already possess. But there clearly are cases where we do acquire concepts by using representations or concepts that we do already possess. For example, we rely on various features of our conceptual repertoire to learn scientific theories through which we acquire our concepts of various unobservable entities. So we need an account of the role mental representations play in mediating the acquisition of new concepts. The view I will develop is a modification and extension of Margolis and Laurence’s sustaining mechanism proposal that I discussed in chapter 4. On the view I will offer, representations figure in concept acquisition in ways analogous to reference-fixing descriptions. This view will exhibit the kind of mediation I set out in the previous chapter and the kind of mediation I will discuss here as different species of a more general common kind of process. Conceptual mediation is one mechanism through which concepts can be acquired.

6.2 Conceptual and Nonconceptual Mental Representations

In chapter 1 I set out some conditions that distinguish concepts from other sorts of mental representations. In the previous chapter I showed how a cognizer might acquire some new mental representations. In this section I will explain the underlying source of these distinctive conditions, and show how mental representations might become conceptualized.
6.2.1 Conditions on Conceptual Representation

To this end, I begin by revisiting the conditions that I listed in chapter 1 as distinguishing specifically conceptual representations:

- **flexibility**: concepts are representations that can be deployed in a wide variety of contexts, to a wide variety of ends and purposes, in a wide variety of projects. In contrast, some conceptions of certain kinds of representations view them as constrained in their application to very restricted contexts. Some mental representations may have only domain-specific application. In contrast, conceptual representation is not limited in this way; it is a hallmark of a concept-using being that it can use the concepts that it has in a wide variety of situations.\(^\text{74}\)

- **access**: concepts are representations that are available to conscious thought and reflection, while other representations may not be. For example, low-level perceptual and linguistic representations are generally thought to be isolated from conscious access.

- **control**: concepts seem to be the kinds of things that are under the control of the organism, while other representations may not be. This is to say, roughly, that an organism can produce tokens of a conceptual representation more or less when it wants to. In contrast, certain other kinds of representations may be tokened only in the context of certain kinds of perceptual stimulation.

\(^{74}\) Notice that this notion of flexibility is not the same as the notion to which I appealed in the previous chapter to explain when appeal to representations is appropriate. That notion of flexibility applies to behavior, where as the notion I have in mind here concerns the ways in which inner representations can be used.
The question I raise now is: do these features tend to co-occur, and if so, why? Do representations that have these features constitute an interesting theoretical kind? Is it a coincidence that these features occur together, or is there some underlying reason? I will argue that there is. The fact that some representations have these features is explained by the system that has them having the capacity to represent its having those representations.

Some preliminary reason to think that these features non-accidentally co-occur comes from reflection on the relations between these different features. For a cognizer to have control over a representation is for it to be able to spontaneously produce tokens of it in thought. But the ability to produce tokens of such a representation at will would seem to require the cognizer to have the ability to be aware of having such token representations, at least sometimes; otherwise it is not entirely clear what sense there is to be made of the idea that productions of the tokens are under the control of the cognizer, and at the very least such an ability would seem to have little point absent the ability to be aware of what one was doing. Similarly, it might be thought, if a cognizer has control over the production of tokens of a given representation, then it can produce those tokens in a variety of situations; otherwise, if the situations in which the tokens could be produced were very restrictive, then that would seem to call into question whether the cognizer really did control the productions of tokens of that representation. So, it would seem, the fact that these features tend to be features of the same mental representations is not simply accidental.

So there is reason to think that the features I have identified as distinctive of conceptual representation are related to each other. Furthermore, there is reason to think
that the co-occurrence of these features is itself explained by some more fundamental underlying feature of conceptual representation. The claim that control, access and flexibility are distinctive of conceptual representation is reminiscent of Gareth Evans’ (1982) Generality Constraint on thought. A cognizer satisfies the Generality Constraint when she can freely recombine her representations: when, for any range of subjects \(a_1-a_n\) and range of predicates \(F_1 \ldots F_n\) that she can represent, she is capable of representing \(F_1a_1 \ldots F_na_1, F_1a_2 \ldots F_na_2, \ldots F_1a_n \ldots F_na_n\). Evans claims that satisfying the Generality Constraint is a precondition of genuine thought; unless a subject can think a variety of thoughts involving \(a\) and a variety of thoughts involving \(F\), she cannot think \(Fa\).

This suggests that satisfying the Generality Constraint is a distinctive feature of conceptual representation. For one thing, Evans himself clearly has in mind that the Generality Constraint should apply to the components of thought, and the conception of thought with which he is concerned is of a piece with central propositional attitudes such as belief, the objects of which are composed of concepts. Furthermore, if a representation satisfies the Generality Constraint, then it also satisfies the conditions I set out above. A representation that satisfies the Generality Constraint will count as flexible, since for a representation to satisfy the Generality Constraint is for it to be combinable with a wide range of other representations the cognizer possesses, while for a representation to be flexible is for the cognizer to be able to use it in a wide variety of different contexts. So, if flexibility is related to control and access in the way I have suggested above, then a representation that satisfies the Generality Constraint will also exhibit the features I have identified as distinctive of conceptual representation.
But what then enables a representation to satisfy the Generality Constraint? What allows it to be combined with a variety of other representations? I suggest that, in order for a system to use a representation in this way, it needs to be able to treat that representation as itself an object of thought. The system that has that representation needs, not simply to be able to use it to manipulate other objects, but to treat it itself as a manipulable object, to be able to perform operations on it. But, in order for a system to treat one of its own representations as itself an object of manipulation, it will have to be able to form a representation of that representation. This means that the system will have to have the general capacity to represent its own representations. So the capacity for metarepresentation is a condition on something’s having conceptual representations.\(^{75}\)

Why should the capacity for metarepresentation and the capacity for conceptual representation be linked in this way? As a way of answering this question, consider a question that Stich (1978) raises about the features of propositional attitudes. Stich there distinguishes between beliefs and what he calls subdoxastic states. Subdoxastic states are cognitive states, but they differ from beliefs in two ways. First, subdoxastic states are not states that the cognitive system that is in them is normally consciously aware of being in, while beliefs are the kind of state that a cognitive system in them is or at least can be consciously aware of being in. Second, subdoxastic states are inferentially isolated, while beliefs are inferentially promiscuous. This is to say that subdoxastic states stand in relatively few inferential relations to other cognitive states, while beliefs can stand in arbitrarily many inferential relations to other cognitive states. The question Stich raises

\(^{75}\) Clark and Karmiloff-Smith (1993) develop a related conception of the importance for genuine thought on a system’s being able to take its own representations as objects for itself, and try to show how such a capacity might be implemented on a connectionist architecture.
is: why are the very same things that are not consciously accessible also inferentially isolated, while the things that are consciously accessible also inferentially promiscuous? What explains why these two features are connected?

I think the answer to this question lies in considering what the point is of a system’s having the ability to be aware of its own inner states. Lycan (1996) argues that the value of this lies in the capacity of such a system to be able to exert greater control over those states, and in being better able to integrate its representations with each other. Awareness of one’s own states allows one to use them together in inference, to eliminate contradictions in thought, to plan, and to reason conditionally and hypothetically. So being consciously aware of one’s own states allows them to be inferentially integrated with each other, and thereby to be inferentially promiscuous. So it is no accident that these two features go together; awareness of one’s own cognitive states allows one to use those states in operations of inference.

Lycan’s own view of conscious awareness is a higher-order perception view. On this view, a cognitive system is consciously aware of being in a certain cognitive state just when there is a state of an inner perceptual system that represents that first-order state. Thus, on this view, the capacity for conscious awareness requires the capacity for metarepresentation. Therefore the capacity for metarepresentation underwrites the increased control that the system has over its own states, and therefore increases the flexibility that the system has in using those states.

So, on this view of what awareness is, awareness of one’s own inner states is a matter of representing those inner states, and the value of the ability to represent those inner states is the increased control and flexibility with respect to those states that such
awareness provides. I think parallel points explain the distinctive features of conceptual representation. Just as Stich claims that beliefs are states of which we can be consciously aware, so I have claimed that concepts, in contrast to nonconceptual representations, are representations of which we can be consciously aware. And while the other feature Stich identifies as distinguishing beliefs from subdoxastic states – that between inferential promiscuity and inferential isolation – has no direct analogue in the distinguishing features I have identified, the reasons that motivate Stich’s distinction are similar to the reasons I have given. Why think that states that are inferentially isolated from other states are not beliefs? Intuitively, because beliefs are states that the system can use in a wide variety of ways, according to the particular purposes it has. Limitations on the inferential relations that a state can enter into are limitations on the flexibility that state has, and the control that the system has over that state.

So the distinction Stich makes between beliefs and subdoxastic states parallels the distinction I have proposed between conceptual and nonconceptual representations. Noting these parallels allows us to offer an explanation of the basis of the distinguishing features of conceptual representation on the model of Lycan’s answer to Stich’s question. Lycan’s solution to Stich’s problem is that conscious awareness and inferential promiscuity go together because the metarepresentational processes that are constitutive of conscious awareness are conducive to increasing the system’s inferential control over those states. Similarly in the present case: control, access, and flexibility go together because the metarepresentational processes that are constitutive of access aid in increasing the control and flexibility that the system has over the representations to which these processes give it access.
So the thought is that the distinctive features of conceptual representation—control, access, and flexibility—are conferred on representations by their being the object of metarepresentational processes. The ability of a system to represent its own representations gives it a certain flexibility with respect to those first-order representations. When a system can treat its own representations as objects, it is able to evaluate them and act on them in ways that it cannot when it merely has the capacity to use its own representations to act on the world. The introduction of this capacity introduces the flexibility that licenses us to say that a system has conceptual capacities.

If flexibility is increased by the process of making a representation the object of metarepresentational processes, then we also have an explanation of why conceptual representations satisfy the Generality Constraint. Explicit representations are freely recombinable with other explicit representations. So when a system develops several different explicit representations, it will be able to recombine those representations with each other, thus satisfying the Generality Constraint.

Whether or not something counts as a conceptual as opposed to a nonconceptual representation is not strictly a matter of the intrinsic features of that representation itself, but instead depends on features of the system that has those representations. A representation-using system becomes a conceptual system, not by changing anything about the representations themselves directly, but by adding additional capacities to manipulate its representations to the system that has those representations.
6.2.2 Conceptual Representation and Conceptual Role

The view of concepts I have proposed is similar in some respects to a conceptual-role theory of concepts, such as that of Block (1986), which I discussed critically in chapter 4. Like Block, I hold that something’s being a concept depends on its relations to other concepts, since, I have accepted the Generality Constraint as a condition on conceptual representation, and satisfying the Generality Constraint requires that the concept be able to be exercised in conjunction with a variety of other concepts.

The central difference between the view I offer here and Block’s view is that, on the latter, the particular conceptual role that a particular concept has at least partly constitutes the identity conditions of that concepts. What makes that concept the very concept that it is, is the fact that it has the particular conceptual role that it does. Block’s view is a theory about (one aspect of) conceptual content.

In contrast, the view I am proposing here is purely a view about what distinguishes conceptual from nonconceptual mental representation. On my view, what makes something a conceptual representation is the fact that it has the features I have described. This may require that a properly conceptual representation has some conceptual role or other, but it does not require that any particular concept have any particular conceptual role. What makes a mental representation a conceptual representation is its possession of the features I have outlined; but what makes it the very representation that it is has to do with its representational content, which is purely a matter of its connection to the relevant feature of the world, not its conceptual role.

So, on my view, there is a holistic aspect to conceptual representation, since what makes a representation conceptual are the relations that it bears to others. But this holism
is not a holism of content, but rather a holism of conceptual status. Because of this, my view is not threatened by the problems that Fodor and LePore (1992) raise for holistic theories of content. According to Fodor and LePore, holism about content is incompatible with the publicity of content, since, if the content of one’s concepts depends on the particular relations they bear to each other, then every difference in these relations constitutes a difference in content. So, for example, if I think cats make good pets and you do not, then this difference in how our concepts are related to each other means that our CAT concepts actually have different contents. But this, Fodor and LePore allege, is incompatible with the generality of psychological explanation, which requires that people’s concepts have the same content.

I will not comment here on the cogency of these concerns about holism. I simply note that, whether or not they are problems for holistic views of content, they do not present any problems for the idea that relations between representations are constitutive of conceptual status. That one person’s CAT representation is conceptual, while another’s is not, is quite compatible with those representations both having the same content.

### 6.2.3 How Representations Become Conceptualized

I have given an account of the distinctive features of conceptual representation. The question now is about how representations could acquire these features. I argue that an extension of the view given in the previous chapter, about the acquisition of new representational capacities, can explain how representations become conceptual.

Recall that, according to the view I have offered here, for a representation to be conceptual is for it to be the object of metarepresentational processes; that is, for the
system to be able to form representations of those representations. Now, in the previous chapter I have shown how to acquire new representation-types. So that story can simply be extended to the acquisition of representations of representations; there is nothing in that account that essentially depends on what the objects of the representations thus formed are. All that this account requires is the presence of some appropriate analogue of perceptual processes that can indicate the presence of representational states, and mechanisms for recording and sorting those states. This role can be attributed to introspection.

So the formation of conceptual representations is not a process different in kind from the formation of new representations more generally. Conceptual representations are formed when the system develops the capacity to represent its own representations. The conditions that are required for the system to exercise this ability are those that distinguish specifically conceptual representations from other sorts of mental representations.

I emphasize, however, that nothing in the results of the previous chapter depend on this particular way of distinguishing conceptual from nonconceptual representations, nor indeed do they depend on any such distinction at all. I have offered this account as a way of drawing the distinction because it explains the phenomena that I identified as distinctive of concepts, and because it is a natural extension of the proposal I offered in the previous chapter. But that proposal is compatible with other ways of distinguishing between concepts and nonconceptual representations. So whether or not one accepts this way of drawing it, one can still accept my account of how mental representations are acquired.
6.3 Concept Acquisition Via Conceptual Mediation

In the previous chapter I showed how concepts can be acquired through perception, without requiring the cognizer to have any mental representations prior to acquiring those concepts. I further presented some empirical reasons to think that this view is plausible in the case of early concept acquisition.

But certainly this is not always how concepts are acquired. We continue to acquire concepts well into adulthood, and many concepts we acquire in adulthood plausibly are acquired at least in part by mobilizing other concepts we have. To acquire concepts of the unobservable postulates of scientific theories, for example, requires learning the theory, which presumably requires having some other conceptual resources already in place.

So concept acquisition by conceptual mediation is a genuine phenomenon. On the other hand, I have already rejected traditional views of concept acquisition as construction from more primitive concepts. So we need another model for the acquisition of concepts on the basis of other concepts.

The acquisition of concepts by means of other concepts presents no problem of principle. In the previous chapter I argued that concepts can be acquired without using other concepts, not that they must be so acquired. And in fact I will argue that in each case the mechanisms of acquisition are fundamentally the same. Other concepts can aid in the acquisition of new concepts; when they do, the role they play is a role that is sometimes played by nonconceptual / nonrepresentational structures. But there are important advantages to using concepts to acquire other concepts. These advantages
explain why, as a matter of fact, some concepts are only ever acquired via the mediation of other concepts.

The traditional view of how concepts mediate the acquisition of new concepts is that the new concept is assembled out of the concepts one already has. On this view, concepts one already has mediate the acquisition of new concepts by being parts of those new concepts. But this idea is problematic, for reasons that I explored in chapter 1: models of conceptual construction appear to be either psychologically implausible or inconsistent with the requirement that concepts be compositional.

As I pointed out there, however, the idea that concepts are used in processes of acquiring new concepts does not entail that they are parts of the products of those processes. That leaves open the possibility that concepts could play a role in the acquisition of unstructured concepts. This is the possibility exploited by Margolis and Laurence’s view of acquiring concepts as acquiring sustaining mechanisms, which I discussed in chapter 4. Here I offer a revision and extension of that proposal.

I propose that we should think of the role of concepts in acquiring new concepts as analogous to the role of reference-fixing descriptions. In his (1972), Kripke distinguished two uses of definite description. A definite description may be used to give the sense of another expression, by providing an expression that is synonymous with it. But it may also be used to indicate or pick out some object, for which a new term can then be introduced. In such a case, the description gives the reference of that term, without being synonymous with it. The description is simply used as a convenient tool for identifying the relevant entity.
The difference between the two uses has to do with whether or not the term and the description have the same modal properties. Where the term and the description are synonyms, then they will not only have the same extension in the actual world, but will also pick out the same things in all possible worlds. Where the description is used just to fix the reference of the expression, on the other hand, the description and the introduction introduced on the basis of it will have the same extension in the actual world, but will diverge in possible worlds distinct from this one. Even though both the description and the name identify the same thing in the actual world, the name is a rigid designator while the description is not, so the description will identify different things in different possible worlds, while the name will identify the same thing across possible worlds. So it is possible to wonder whether the thing named might have failed to satisfy the description, even when the name is originally introduced by means of a description that fixes the referent of the name. Since meaning is, at least, a function from possible worlds to extensions, this is a way of introducing a term with a new meaning.

Kripke discusses reference-fixing descriptions in the context of discussing names and singular terms. But, as Putnam (1975) shows, similar morals apply to the case of general terms. Putnam’s major point there is that the meaning of many of our terms are not analytic constructions out of other terms, but depend on the environment we inhabit. But, Putnam claims, competence with these terms still requires having some ways of identifying members of the extension of the term. Being competent with the term “tiger” may, for example, require knowing the “tiger stereotype”: knowing that tigers are large striped cats, for example.76

---

76 I am not here carefully distinguishing between the introduction of a term into a public language and the introduction of a term (which may already be a part of a public language) into a speaker’s idiolect, which is
It may be too much to require, as Putnam does, that there is some particular body of information that one must have in order to be competent with a term such as tiger. A weaker requirement would be that one should have some bit of information or other about the extension of the term in order to count as competent with it. Alternatively, one might reject the idea that one counts simply as either competent or not with a term, and opt instead for a graded notion of competence, according to which a person counts as more competent with a term the more information he has about the extension of that term. However this may be, the important point for present purposes is the parallel with Kripke’s position on the introduction of singular terms. Just as singular terms can be introduced by the use of descriptive information to identify the intended referent, so general terms can be introduced by appeal to descriptive information to identify the intended referent as well.

Kripke’s and Putnam’s views are in the first instance views about language and competence with linguistic items. But analogous considerations apply in the case of thought. A description in thought can be used to identify a particular item or category, for which a new mental term can then be introduced. In such a case the description is used to fix the reference of this term, but not to give the sense of it. The description does not give the conditions that something must satisfy in order for the term thus introduced to apply to it; that is, it is not a specification of the satisfaction-conditions of the mental term thus introduced. It is merely a way of identifying something so that it can serve as the referent for a new term.

what I have in mind when I speak of a speaker’s acquiring competence with a term. This is because, in the analogous case of concept acquisition, introducing a new mental term is always a matter of introducing it into a speaker’s idiolect. So the difference between the two in the case of natural language will not be relevant for my purposes.
In the case of the reference-fixing use of descriptions, the description is just used to convey the relevant audience to the right individual. The description is thus fundamentally analogous to a kind of demonstration. Descriptions, in their reference-fixing use, and demonstrations are each used to identify some object or feature of the environment. The reference of a new term may be fixed either directly, as in the case of a demonstration, or indirectly, as when a description is used, but the relation between the term thus introduced and the referent thus identified is the same in either case.\textsuperscript{77}

One other feature of the use of descriptions to introduce new terms that bears mentioning is highlighted by Donnellan’s (1966) distinction between the attributive and referential use of definite descriptions. According to Donnellan, we can use definite descriptions either to characterize the satisfier of the description, whatever it is, or as a means to identify some individual who we believe to satisfy the description. When we use descriptions referentially, we are using them in a way analogous to our use of demonstrations. What Donnellan’s discussion highlights is that, when a description is used referentially, whether or not it in fact accurately describes the intended referent may be beside the point. If I say to you “The man in the corner drinking a martini is kind”, my description may succeed in identifying the relevant person even if he is drinking water out of a martini glass. Suppose then we introduced a name on the basis of this identification; intuitively the name would refer to the man identified, whether or not he in fact satisfies the description. So a description may serve to introduce a term into a

\textsuperscript{77} “Indirectly” here means that a computation of the semantic value of the description is required to effect the identification. The description will be composed of concepts, and so understanding the description will require that one have the concepts that figure in it, which understanding will be independent of its role in identification (one can understand a description without knowing what if anything satisfies it). In contrast, a demonstration is not a semantically structured item in the relevant sense, so the relation between it and its referent is unmediated by any prior determination of semantic value.
language even when the description does not accurately describe the individual in question. In analogous fashion, false beliefs and other incorrect descriptive information may still be sufficient to identify the right referent for a mental term. So, importantly, that descriptive information can mediate the acquisition of concepts is compatible with the idea that we can have concepts despite having false beliefs about the objects of those concepts, and that we can share concepts despite having differing beliefs.

I have been arguing that we can understand the role that concepts play in mediating the acquisition of new concepts as analogous to the role that descriptive information plays in introducing new terms into our competence with a natural language. I have also argued that, in this role, descriptions function just as do demonstrations.

So, from this perspective, descriptions and demonstrations can be seen as different mechanisms for a common process. The significance of this is that, as I argued in the previous chapter, concepts can be introduced by a process akin to demonstration. The account of concept acquisition that I offered in the previous chapter is fundamentally a demonstrative kind of term introduction, relying on nonrepresentational perceptual states to effect the identifications which allow the introduction of new representations. Since descriptions can be used in much the same way as demonstratives, and the use of concepts to acquire other concepts should be understood on the model of this use of descriptions, the use of concepts to acquire other concepts is a more sophisticated way of performing precisely the same task as is performed by perceptual demonstration. This way of describing something is, in the relevant respects, a complicated way of pointing to it.  

78 Sholl and Leslie (1999) and Ballard (1991) defend and develop views of perceptual object tracking that suggest a demonstrative model of such processes.
This model of the role that concepts play in mediating the acquisition of new concepts, at the same time it explains what acquisition by representational mediation has in common with acquisition without representational mediation, also explains at once why concept acquisition via perception is the central case of early concept acquisition, and what the point of using concepts to acquire other concepts is. Identification in perception is limited by one’s own spatiotemporal position. In order for a demonstrative identification to be successful, one must be able to locate the object demonstrated in egocentric relation to oneself. Objects that are out of one’s perceptual range, and that cannot therefore be ostended, cannot have representations introduced via demonstration. With the ability to use descriptions, however, this constraint is loosened. One can describe someone as “the man in the corner drinking a martini” and succeed in identifying someone even when the individual is currently out of perceptual range. And, while that is an example of a description that itself contains egocentric elements, descriptions can also be formulated in purely impersonal terms.

So the use of descriptions allows us to introduce terms into our representational systems that are spatiotemporally removed from us. As we acquire more conceptual resources, then, we are better able to identify things that stand out of our immediate perceptual reach, whether this be because they are too far away from us to perceive – as in the case of planets in other galaxies – or because they are imperceptible to our perceptual organs – as in the case of germs or electrons. So the point of using concepts to acquire other concepts is to extend the range of those things for which we can acquire concepts beyond what we can identify by means of perception.\footnote{Millikan (2000) offers a similar picture of how the tracking of aspects of our environment by means of concepts is an extension of our tracking aspects of our environment through perception, although I do not}
This point also explains why some concepts we have are only, as a matter of fact, only acquired through the mediation of other concepts that we have. Since germs do not, as a matter of fact, make any impression on our sensory systems, we need further resources to introduce such concepts into our repertoires. The reason for this is essentially the same as the reason why we (in North America) need concepts to introduce **koala** into our conceptual repertoires – because there are no koalas around here to look at.

The question remains whether there is any evidence to support this model of how concepts are acquired by means of other concepts. In fact I think there is a great deal of evidence for this model, once it is looked at in the right way. Specifically, I think the literature on the reality of prototype effects (e.g., Rosch et. al. 1976; Smith and Medin 1981; Prinz 2002) provides evidence for this model of conceptual mediation. The philosophical literature on prototypes has tended to revolve around the question of whether prototypes could be concepts, which in turn has centered around the question of whether prototype representations exhibit important semantic features of concepts such as compositionality. In my view, this is a misunderstanding of the role of prototypes in a theory of concepts.\(^{80}\) Prototypes are miscast as theories of conceptual content, but they are well-placed to explain concept acquisition nonetheless, for we can think of them as

\[^{80}\] Strangely Laurence and Margolis do not discuss prototype theories of concepts at all, despite their apparent affinity with their sustaining mechanism view; in particular, what they identify as “syndrome-based” sustaining mechanisms seem to be very much like prototypes. It is possible that this is because think of prototype theory as a theory of conceptual content. But, as I have suggested, this is the wrong way to think about prototype theories of concepts.
the analogues of reference-fixing descriptions. Like descriptions of categories, prototypes contain typical, salient, and diagnostic information about the categories to which they refer. And, like descriptions, prototypes may contain inaccurate information, so long as the information they do contain identifies the appropriate category. So prototype theory does provide the seeds of a plausible account of how at least some concepts are acquired, once it is recognized that such theories do not seek to provide an account of the content of concepts, nor are they committed to the claim that concepts thus acquired are ipso facto semantically structured.

So concepts can be acquired by means of other concepts in much the same way that new terms can be introduced into a language by means of reference-fixing descriptions. This method of introduction does not require that the concepts thus introduced be structured, and thus can provide an account of the acquisition of primitive concepts. Furthermore, since the role for descriptions to which it appeals is essentially the same role as the role I characterized perception as playing – in each case, the job is to identify something so that a representation can be introduced for it – there is no explanatory asymmetry between how concepts are acquired by means of other concepts and how concepts are acquired without reliance on other concepts. This means that we have a unified theory of concept acquisition, one that exhibits the different mechanisms that lead to concept acquisition as different ways of implementing a common process. (This is why, for example, those of us in North American acquire our KOALA concept by reliance on other conceptual resources, while those in Australia acquire their KOALA concept without needing to rely on any other conceptual resources, but what we acquire in each case is the same.)
6.4 Conclusion

In this chapter I have accomplished two things. First, I have explained the difference between conceptual and nonconceptual mental representations, and how the distinctive features of conceptual representation can be acquired. In doing so I have closed a gap in my account of acquisition from the previous chapter, which showed how to acquire a new representational structure, but was silent on what more there might be to conceptual representation than simply mental representation more generally. The account in this chapter completes my account of concept acquisition.

Second, I have explained how concepts one already possesses can mediate the acquisition of new concepts, in a way compatible with my account of concept acquisition. On this view, conceptual mediation is one mechanism for acquiring new concepts, one with distinctive advantages, but one that is another way of performing the same function that can also be performed without benefit of conceptual mediation. The account in this chapter underscores the generality of my account of concept acquisition. By relying on the various mechanisms we have, both those we start out with and those we acquire in the course of development, we can begin from an initial state containing no mental representations at all, and progress to the acquisition of a complete conceptual repertoire.

There are many ways in which the view I have offered in these chapters is incomplete. It gives, for example, no account of the acquisition of logical/mathematical concepts, moral and evaluative concepts, or concepts pertaining to the “theory of mind”, for example. While I believe satisfactory explanations of the origins of these concepts can be given within the framework I have provided here, for the present they must be left
as topics for further research. I can claim, however, that the account I have offered provides a promising framework for investigating these questions, since it shows us how to build a large portion of our conceptual repertoire from scratch.
REFERENCES


