

UCLA

UCLA Electronic Theses and Dissertations

Title

Forming Concepts and Doing it Well: A Theory and Epistemology for Empirical Concepts

Permalink

<https://escholarship.org/uc/item/8ss7n0q4>

Author

Weber, Tamar

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Forming Concepts and Doing it Well:
A Theory and Epistemology for Empirical Concepts

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Philosophy

by

Tamar Weber

2017

© Copyright by

Tamar Weber

2017

ABSTRACT OF THE DISSERTATION

Forming Concepts and Doing it Well:

A Theory and Epistemology for Empirical Concepts

by

Tamar Weber

Doctor of Philosophy in Philosophy

University of California, Los Angeles, 2017

Professor Tyler Burge, Chair

In this dissertation, I argue that we need an epistemology of empirical concept formation in order to explain the epistemic continuity from perceiving the world to achieving empirical knowledge.

I situate empirical concept formation within the greater epistemic endeavor I call “empirical discovery.” The aim of empirical discovery is to form warranted empirical beliefs—and ultimately—to achieve knowledge and understanding of the nature and structure of the empirical world. The challenge of discovery is that to learn and understand empirical truths is not merely to compose and confirm a list of hypotheses and attendant theories. Discovery requires creating the representational tools from which such hypotheses and theories are composed. Empirical concepts are some of these representational tools. Thus, forming empirical concepts is part of the greater epistemic endeavor of empirical discovery.

My account explains how forming concepts on the basis of perceptual experience can yield prima facie warrant to employ such concepts, even in cases where one knows nearly nothing about the properties indicated by such concepts. I call that prima facie warrant “concept formation

entitlement.” According to my theory, our warrant for forming beliefs about empirical subject matter of which we know nearly nothing, is inextricably linked to the process by which we form new ways of representing the world.

Drawing on research in cognitive, developmental, and perceptual psychology, I offer an empirically plausible model of how we form a wide range of general empirical concepts. I explain the connection between the model and the fulfillment of epistemic norms that govern empirical belief revision on route to discovery.

As a prelude to my positive proposal, I trace Jerry Fodor’s path to what I call his “new radical conclusion.” His new radical conclusion is that concept acquisition is not a cognitive process, does not admit of psychological explanation, and cannot be explained at the intentional level. I reconstruct six variants of Fodor’s argument, and argue that all of them fail. I use the epistemic underdetermination variant of his argument as a springboard for the thesis that concept acquisition is epistemically evaluable.

The dissertation of Tamar Weber is approved.

Samuel Cumming

Mark Greenberg

Edward Stabler

Tyler Burge, Committee Chair

University of California, Los Angeles

2017

For my parents,

Mira and Eli

Contents

Introduction	1
0.1 Two guiding questions.....	1
0.2 A note on terminology.....	2
0.3 What are GEM concepts?.....	3
0.4 The origins of the guiding questions.....	3
0.4.1 Question 1: How can new ways of representing be created?	3
0.4.2 Question 2: Are concept formation processes subject to epistemic evaluation?.....	5
0.5 Chapter summaries.....	7
0.5.1 Chapter 1: Fodor’s New Radical Conclusion.....	7
0.5.2 Chapter 2: How to Acquire a Concept Well (epistemically speaking).....	9
0.5.3 Chapter 3: Concept Formation Entitlement and Belief Revision.....	9
1 Fodor’s New Radical Conclusion	12
Introduction.....	12
1.1 Fodor Fundamentals.....	13
1.1.1 Psychological/intentional realism.....	13
1.1.2 Naturalism.....	13
1.1.3 The representational and computational theory of mind.....	13
1.2. The Arguments of <i>LOT1</i> and <i>LOT2</i>	16
1.3 Fodor’s new radical conclusion.....	21
1.3.1.Fodor’s positive proposal.....	22
1.3.2 A closer look at Fodorian stereotypes.....	23
1.3.3 A closer look at P2: acquiring a concept as a consequence of learning a stereotype.....	25
1.4 Fodor’s path to the new radical conclusion.....	26
1.4.1 Fodor’s argument in its most general form.....	26
1.4.2 The Projection Problem and NRC(c)-(f).....	28
1.4.3 Epistemic underdetermination, the argument from arbitrariness, and Fodor’s brute/rational dichotomy.....	31
1.4.4 Semantic underdetermination and computational processes.....	33
1.4.5 Fodor on computational processes more generally.....	36

1.5 Conclusions and reflections.....	39
2 How to Acquire a Concept Well (Epistemically Speaking).....	43
Introduction.....	43
2.1 Motivating an affirmative answer to the evaluability question.....	46
2.1.1 Forming natural kind concepts in pursuit of knowledge.....	46
2.2 The epistemic contribution of concept formation processes.....	51
2.3 A model for acquiring natural kind concepts.....	54
2.3.1 The Essentialist Model.....	54
2.3.2 The Projection Problem revisited: stereotypes underdetermine concepts.....	57
2.3.3 A solution to The Projection Problem.....	58
2.4 The Epistemic Challenge applied to The Essentialist Model.....	60
2.4.1 The Epistemic Challenge reintroduced.....	61
2.4.2 An evidential condition.....	62
2.4.3 A reliability condition.....	66
2.4.4 Representational function and reliability.....	72
2.5 Concept formation entitlement.....	76
2.5.1 Source1: The nature and function of our sense-perceptual capacities.....	78
2.5.2 Source2: Making a contribution to the epistemic norms of empirical discovery.....	80
2.6 An objection to the need for concept formation entitlement.....	82
2.7 Conclusion.....	84
3 Concept Formation Entitlement and Belief Revision.....	86
Introduction.....	86
3.1 Part I: What are GEM concepts?.....	88
3.1.1 General, Empirical, Mediumish.....	88
3.1.2 Kind sortal GEM concepts.....	90
3.1.3 The role of perception in GEM concept formation.....	91
3.2 Part II: How to form a GEM concept, the empirical model.....	92
3.2.1 A model of GEM concept formation and stereotype creation.....	92
3.2.1.1 Stage 1: Forming a perceptual state and a basic perceptual belief.....	93
3.2.1.2 Stage 2a: Selecting a concept formation path.....	94

Exhibiting a biological motion pattern.....	96
Self-initiated motion.....	97
Responding contingently at a distance.....	98
3.2.1.2b Stage 2b: Selecting a perceptual feature space.....	98
3.2.1.3 Stage 3: Forming A GEM Concept.....	101
3.2.1.4 Stage 4: Creating an associated stereotype.....	103
3.2.2 The model in full.....	105
3.3 Part III: The place of concept formation in the epistemology of empirical discovery.....	107
3.3.1 The defeasible presumptions that back GEM concept formation and early application.....	109
3.3.1.1 The Veridical Perception Presumption.....	109
3.3.1.2 The No Global Fakes Presumption.....	111
3.3.1.3 The One-Look-One-Kind Presumption.....	113
3.3.1.4 The Fakes-And-Deviants-Are-Rare Presumption.....	114
3.3.2 The role of the defeasible presumptions in concept formation entitlement.....	116
3.3.2.1 The defeasible presumptions reliably track truth.....	117
3.3.2.2 The defeasible presumptions secure and set the limits of warranted belief revision.....	119
3.4 Part IV: Reflections and conclusions.....	121
Conclusion.....	125
Bibliography.....	136

List of Figures

3.1 GEM Concept Formation and Associated Stereotype Creation Process	93
3.2 Fifteen Particulars	100
3.3 A Two-Dimensional Feature Space	100
3.4 From Perception to Concept Formation and Associated Stereotype Creation . . .	106

Acknowledgements

First and foremost, I would like to thank my committee members. Tyler Burge's encouragement, guidance, and scrutiny paved the path for this dissertation. His exceptional body of work continues to intellectually impress and challenge me at every turn. The Burgean foundation of this dissertation merely scratches the surface of my philosophical appreciation for the depth of his ideas. Sam Cumming's continual encouragement, both philosophically and otherwise, was more helpful than he could know. Mark Greenberg has been an invaluable philosophical interlocutor. His own work, seminars, and feedback made a tremendous impact on my thinking about the issues at the heart of this dissertation. Ed Stabler was a wonderful resource in early days when I was learning about learnability. His support and outsider perspective have been a great help. Many thanks to you all.

For valuable conversations, both in print and in person, I thank Ned Block, Susan Carey, Eric Margolis, Christopher Peacocke, and Michael Rescorla.

For philosophical stimulation and/or friendship over too many years, I must thank Denis Beuhler, David Friedell, Marina Folescu, Mikkel Gerken, Jonathan Gingerich, Gabe Greenberg, Andrew Jewell, Gabby Johnson, Kevin Lande, Hakwan Lau, S.M. Love, Megan Metcalf, Eliot Michaelson, Eileen Nutting, Gabe Rabin, and Alex Radulescu.

I owe special thanks to Laura Gillespie, Rachel Johnson, and John Brown. Each of you, in your own special way, provided a respite from the daily grind of graduate school. And to Katie Elliott and Ashley Feinsinger, your humor and spirit sustained me during countless hours of writing, revising, and head-in-hands-grumbling.

Most of all, thank you to my family. Mom and Dad, your boundless love, ceaseless encouragement, and unwavering support make the impossible seem possible. To my sister, Michal, I thank you for your friendship and the strength you channeled my way over many difficult years. And a special thank you to Sherwood Hachtman. I could not have made it through this last year without you.

Tamar Weber

Vita

Education

C.Phil, Philosophy, University of California Los Angeles, 2011

M.A., Philosophy, University of California Los Angeles, 2009

M.A., Political Theory and Political Philosophy, New York University, 2004

B.A., *Magna Cum Laude*, Philosophy (major), English and American Literature (minor)

New York University, 2003

Publications

Weber, T. and Lau, H. (2015). “Consciousness, Neural Basis of,” *International Encyclopedia of Social & Behavioral Science*, 2nd ed., vol. 4, 673-8. Elsevier Science.

Presentations

“The Search for NCCS and Block’s Access/Phenomenal Distinction.”

- Invited guest lecturer for *Psych 88SA Seminar: The Enigma of Conscious Experience*.
UCLA Psychology Department. May 2015

“How to Acquire A Concept Well (Epistemically Speaking)”

- *Society for Philosophy and Psychology*. Vancouver, BC Canada June 2014
- *UCLA Philosophy Department Albritton Society*. Dec 2013

“When Perception Goes Cross-Modal”

- *Towards a Science of Consciousness Conference*. Tucson, Arizona April 2012
- UCLA Philosophy Department Albritton Society. March 2012

“Ignorance, Error, and Belief Revision”

- UCLA Philosophy Department Albritton Society. June 2011

Invited Conference Participation and Comments

- Invited Participant. *40th Oberlin Colloquium in Philosophy*. Oberlin College, May 2012
- Comments on S. James, “De Re Hallucination as Objectual Remembering.”

USC/UCLA Graduate Student Conference. USC, Feb 2012

- Invited Participant. *More or Less: Varieties of Human Cortical Colour Vision*. Vancouver, BC Canada Aug 2011
- Comments on D. Friedell, “Some Metaphysical Issues About Abstract Objects.”

UCLA Albritton Society Conference. May 2011

Introduction

0.1 Two guiding questions

The twin guiding questions of this dissertation are the following:

1. How can new ways of representing that go beyond perceptual representation be created?
2. Are concept formation processes subject to epistemic evaluation?

The second question fits neither squarely in the concept acquisition literature nor in the epistemology literature. The second question grew out of my thinking about the first. The first question demands an explanation of how experience can contribute to concept formation. Thus it focuses attention on the nature of the relationship between experience and concept formation, and what seemed intuitively like a psychological transition with veridicality or correctness conditions. However, concepts cannot be true, false, veridical, or non-veridical; they lack an obvious place within epistemology. Thus the goal of the dissertation became to develop an empirically plausible model of concept formation, and to explore the way in which concept formation can make a positive epistemic contribution to the formation and preservation of true beliefs.

Ultimately I argue that our warrant for forming beliefs about empirical subject matter of which we know nearly nothing is inextricably linked to the process by which we form new ways of representing the world. I argue that there are epistemic standards by which we judge the process of acquiring a concept. Only some cases of acquiring a concept provide prima facie warrant for employing that concept. These cases are the ones in which the formation processes are non-accidental relative to our capacity for veridically representing objective empirical subject matter.

In the remainder of the introduction will proceed as follows: I begin with a terminological clarification, and an explanation of GEM concepts, the class of concepts to which my concept formation model applies. I then detail the intellectual origins of the twin guiding questions, and how they jointly point to the need for an epistemology of concept formation. Finally, I close by summarizing the goal and scope of each of the chapters of the dissertation.

0.2 A note on terminology

Throughout the dissertation I use the expressions ‘concept acquisition’ and ‘concept formation.’ The latter is non-standard in philosophy and (to a less extent) in developmental psychology. I suspect that the literature’s use of acquisition-talk is more likely a reflection of the influence of the language acquisition literature rather than a conscious choice.

I find the expression “concept acquisition” misleading in so far as it primes a picture of concepts being handed over or dropped into the mind as opposed to something that is created by the mind. Acquisition-talk underplays the active role played by psychological processes that create new ways of representing. I prefer to talk in terms of concept formation, forming a concept, etc.

Some might find this to be a minor terminological quibble but I think the change marks a shift from thinking about the outputs of psychological processes to thinking about the production of the outputs. Formation is a species of acquisition in the same sense that baking a cake or sculpting a statue is a way of acquiring a cake or a acquiring a statue respectively. Baking and sculpting are ways of acquiring cakes and statues but they are distinctly *productive* ways of doing so. The expression ‘concept formation’ is meant to highlight the distinctly productive nature of the processes by which we acquire concepts.

In Chapter 1 I use the expressions ‘concept acquisition,’ and ‘acquire a concept’ rather than ‘form a concept’. My reason for doing so is that Chapter 1 is a close reading of Jerry

Fodor's writing and he writes exclusively in terms of acquisition. Chapter 2 is, in many respects, a bridge between the dominant literature on concept acquisition and my own positive proposal. As such I write using both 'acquisition' and 'formation' depending on context. Chapter 3 is devoted entirely to my positive proposal and thus I write entirely in terms of concept formation.

0.3 What are GEM concepts?

The scope of the dissertation is restricted to a wide range of concepts I call "GEM concepts." These concepts include many natural kind concepts such as KOALA, LEMON, and CARIBOU; artifact kind concepts such as CARABINEER, KNISH, and PHONE; as well as many other concepts that defy easy classification such as MOUNTAIN, LAKE, and PUDDLE. The acronym 'GEM' is derived from the three characteristics that unify the paradigmatic concepts that the model covers. GEM concepts are *general* concepts that indicate objective properties that are instantiated in whole particulars that can be picked out in perception. Such properties are the subject matter of *empirical beliefs*. The particulars in which such properties are instantiated tend to be *medium-ish* in size. Mediumish objects are perceptually discriminable independently of theory (though not necessarily independently of some instrument). Given the role of perception and perceptual cues in concept formation and my interest in ontogenetically early empirical concepts, the restricted category is well motivated. I expound the notion of a GEM concept in detail in Chapter 3.

0.4 The origins of the guiding questions

0.4.1 Question 1: How can new ways of representing be created?

My interest in the possibility of creating new ways of representing began after reading a series of papers written by Hilary Putnam and Tyler Burge. (See in particular Putnam 1970/1975c, 1973/1975b, 1975/1975d and Burge 1979/2007a, 1982/2007c, and 1986/2007b). Together, the papers provide an extended examination of the nature of linguistic meaning, the

individuation of mental states, and the notion of incompletely understood concepts. The papers include a multitude of thought experiments that vividly exemplify the fact that we can represent and think about properties about which we have very little knowledge and very few true beliefs. In particular Burge's account of the phenomenon of incomplete conceptual understanding shows that an individual can have the minimal understanding needed for concept possession without understanding the constitutive principles that govern its usage (Burge, 2007d, p.163). That the phenomenon of incomplete understanding abounds was clear. However, what was less clear was a characterization of the ways in which an individual can achieve the minimal understanding needed for concept possession.

My interest in minimal conceptual understanding in conjunction with reading Jerry Fodor's *The Language of Thought* (1975) placed my research squarely in the concept acquisition literature. Though *The Language of Thought* is predominantly a treatise on a computational framework for investigating cognition, in it Fodor articulates a puzzle that is continuous with the anti-descriptivist sentiments of the papers described above. The puzzle begins by accepting the claim that the meaning of a term cannot be given by a definition, and then brings that claim to bear on the possibility of learning a concept. In response Fodor famously argues himself into radical concept nativism—the view that most of our definitionally primitive concepts must be innate. Like many people I read Fodor's argument for radical concept nativism as a compelling paradox rather than a proof. Reading Fodor alongside the papers by Putnam and Burge raised the following question: Once we disabuse ourselves of the possibility of definitions and accept that ubiquity of incomplete understanding, how can we begin to represent the world around us? The question struck me as most pressing in the context of early development because this is where the problem is most stark. Because mature adults have both a large body of knowledge about the

world and the ability to reflect on their own conceptions, the problem of learning a new concept seemed at the very least tractable. But infants and early language learners had, it seemed to me, such impoverished representational abilities compared to adults that the possibility of concept learning was enigmatic.

The question of how we can begin to represent the world cannot be addressed strictly from the armchair. That claim will seem controversial to some philosophers and entirely obvious to others. The question of how we actually begin to represent the world strikes me as far more interesting than a characterization of how we could form concepts in all possible worlds. Accordingly I turned to developmental psychology and later supplemented that literature with perceptual psychology. Susan Gelman's account of psychological essentialism in *The Essential Child* (2003) informed my early thinking about the subject. Psychological essentialism is a theory about the psychological tendency to essentialize categories. The central claim of psychological essentialism is that people (including children as young as 2-3 year olds) distinguish between the superficial properties and the essential unobservable deep structural properties that type-identify objects and are responsible for observable similarities that category members share. Gelman's defense of the claim that several independent capacities individually and jointly lead to the manifestation of psychological essentialism is partly inspired by and naturally complemented Putnam (1975/1975d) and Burge's (1982/2007c) discussions of natural kind concepts and defeasible stereotypes. Examining the results of developmental and perceptual psychology experiments began to shape my thinking about how minimal conceptual understanding can be instantiated.

0.4.2 Question 2: Are concept formation processes subject to epistemic evaluation?

That infants and early language learners form concepts under conditions of massive ignorance underlies Fodor's use of the expression 'concept triggering' in an epistemically

disparaging sense. This seemed wrong given that infants and toddlers regularly form concepts in response to experiences of typical instances of the property indicated by those concepts. In reaction against Fodor's own work I began thinking about concept acquisition as an epistemically evaluable process. Thus I formed the second guiding question of the dissertation and developed it into two questions that I call "the evaluability question" and "the epistemic challenge."

The evaluability question asks: Is concept formation subject to epistemic normative standards at all? The epistemic challenge presupposes an affirmative answer to the evaluability question and asks: Under what conditions is concept formation warranted?

In developing my thinking about warranted concept formation I employed Burge's teleological framework for explicating perceptual entitlement. (Burge, 2003) Entitlement is warrant without reason. It is warrant that derives purely by a psychology's operating well and reliably with respect to its representational functions. Burge offers a teleological account of representational states that contribute to epistemic entitlement to hold perceptual beliefs on the basis of perceptual experience. He draws a connection between the epistemic norms that govern representational states and processes, and the representational function of such states and processes to represent veridically.

I endeavor to use a similar framework in my account of concept formation entitlement. Concept formation entitlement is an epistemic prima facie warrant for empirical beliefs that employ newly formed GEM concepts. However, concept formation entitlement cannot straightforwardly parallel perceptual entitlement. The epistemic norms that Burge describes are the norms governing the transitions from perceptual states to the formation of perceptual beliefs. The transition occurs between two representational kinds that have veridicality conditions. The

concept formation process eventuates in concepts; and as I noted earlier, concepts cannot be true, false, veridical, or non-veridical. Moreover epistemic standards apply to representations and processes that contribute to the formation or preservation of true beliefs. Concepts are not beliefs, so the role of concept formation entitlement in the formation or preservation of true beliefs was unclear.

The location of concept formation entitlement within epistemology became clear as I reflected on the epistemological dimension of Burge's extended discussions of incomplete understanding. The warrant underlying the minimal understanding of newly formed concepts was at least in part explainable in terms of the well functioning of processes which produced the newly formed minimally understood concepts. Thus concept formation entitlement has a role to play in our larger epistemic projects: it is a *prima facie* warrant to employ newly formed minimally understood concepts in beliefs.

0.5 Chapter summaries

0.5.1 Chapter 1: Fodor's New Radical Conclusion

Chapter 1 is a close reading of Jerry Fodor's most recent discussion of concept acquisition as presented in *LOT 2: The Language of Thought revisited* (2008). Beginning with the publication of *The Language of Thought* (1975) Fodor's impact on the concept acquisition literature is unquestionable. An adequate review of the various twists and turns that brought Fodor to his present view would swamp any further issues. Rather than being a comprehensive review, Chapter 1 focuses on what I call "Fodor's new radical conclusion." I formulate the new radical conclusion as follows:

Concept acquisition is not a cognitive process, is not in the domain of cognitive psychology, does not admit of psychological explanation, and cannot be explained at the intentional level.

I argue that, far from being a simple paradox, Fodor's argument for the new radical conclusion is a consequence of several long held commitments that pervade the Fodorian cannon.

In the context of the full dissertation Chapter 1 serves two purposes: First, it introduces what I call the "projection problem," which I draw out of Fodor's discussion of his positive proposal for a concept acquisition model. The projection problem arises because of the underdetermination of concepts by the experiences that elicit them. Specifically, Fodor argues that concepts are underdetermined by representations that he calls "stereotypes." Stereotypes are statistical representations of (purportedly) frequently co-instantiated properties in the world. Fodor uses the underdetermination of concepts by stereotypes as part of his argument for his new radical conclusion.

In the context of the full dissertation Ch.1 serves a second purpose: it introduces the idea of epistemically evaluating concept acquisition processes. My analysis of Fodor's use of the projection problem disentangles a cluster of commitments that drive Fodor to his new conclusion, and in so doing reveals a several different questions that we can ask about the concept acquisition process. For example, we can distinguish the question of whether the process is computational from the question of whether it is a species of rational inference. Fodor frequently runs such questions together. The analysis makes it possible for such questions to be considered independently of one another. Of the six distinguished questions the one that is most central to the dissertation is the question of whether experience can play an evidential role in the process of forming a new concept. While I agree with Fodor that the answer is "no," it suggests a more general and interesting question: Can the process of concept acquisition be epistemically

evaluated? This is one of the guiding questions of the dissertation and the central question of Chapter 2.

0.5.2 Chapter 2: How to Acquire a Concept Well (epistemically speaking)

Chapter 2 is organized around what I call “the evaluability question” and “the epistemic challenge” (described above). I argue that concept formation is part of the epistemic endeavor of empirical discovery. When concepts are formed in an epistemically good way, they yield warrant to employ the newly formed concepts. That warrant is what I call “concept formation entitlement.” I argue that concept formation entitlement fills a gap in the epistemology of empirical beliefs, and thus ultimately in the epistemology of scientific knowledge.

To yield warrant, concepts must be formed in a way that sets up a representational connection between concepts and the properties they indicate in a non-accidentally reliable way. The representational connection must be non-accidentally reliable with respect to the representational capacities of the psychology that has formed the new concept.

On my proposal, the entitlement (and thus warrant) achieved in forming concepts on the basis of perceptually derived representations has two sources: (1) the nature and function of our sense-perceptual capacities by which perceptual states and stereotypes are formed; and (2), the contribution that the process makes to the fulfillment of epistemic norms that govern the formation of empirical beliefs and empirical discovery. I introduce one such set of norms that I call “revisability norms” which I explicate in detail in Chapter 3.

0.5.3 Chapter 3: Concept Formation Entitlement and Belief Revision

Chapter 3 is wholly devoted to my positive proposal for an empirically informed concept formation model and its relation to a normative account of empirical belief revision. My concept formation model applies to GEM concepts (described above). I explain how a concept formation process can epistemically support employing newly formed concepts in our earliest empirical

beliefs. Concept formation entitlement in part explains how we can be warranted in forming such beliefs in the face of tremendous ignorance.

I explain GEM concept formation and stereotype formation in four stages. The first stage is the formation of a perceptual state and a perceptual belief. Perceptual cues present in the perceptual state function as biases that set the concept formation process along one of three paths. Each path corresponds to one of the following three global categories: animate, non-living mover, and self-mover. Examples of animate concepts include animal concepts such as DOG, KOALA, and BIRD. Examples of non-living mover concepts include concepts of artifacts that are frequently seen endogenously caused motion such as PLANE, CAR, and TRAIN. Examples of non-mover concepts include ROCK, SCREWDRIVER, and CARABINEER. Concepts that fall under each of these three categories are formed on the basis of representations of frequently co-instantiated perceptual features such as shape, texture, motion pattern, and coloring, as well as particularly salient higher level features such as eyes, hands, faces, etc. I call these statistical representations “feature clusters.” Once a concept is formed, a stereotype of the property indicated by that concept is associated with the newly formed concept.

The biases described in the empirical model make the problem of forming concepts tractable. However, that same poverty of information makes it difficult to see how using concepts in our earliest beliefs can be warranted. I present a normative account in which each stage of the GEM concept formation model corresponds to one of four defeasible presumptions.

The defeasible presumptions are as follows:

Presumption 1: “Veridical Perception Presumption”

The perceptual system is reliable in the present environment, and it produces perceptual states that are veridical.

Presumption 2: “No Global Fakes Presumption”

There are no cross-global-type fakes.

Presumption 3: “One-Look-One-Kind Presumption”

There is one kind for every feature cluster. An unfamiliar looking particular is an instance of a new kind, and a familiar looking particular is a particular of a previously observed kind.

Presumption 4: “Fakes-And-Deviants-Are-Rare Presumption”

The sample on the basis of which a stereotype is formed is not made up of fakes or deviants.

All of the defeasible presumptions are underwritten by a presumption of normality, that is, the presumption that the present environment is the normal environment. The defeasible presumptions support the formation of warranted early empirical beliefs in two ways. First, the presumptions track real patterns and kinds in the world. Second, the presumptions jointly secure the revisability of most of our empirical beliefs. A belief is empirically revisable if there is an epistemically warranted route to the rejection of the belief, given some possible empirical evidence. Ultimately, the defeasible presumptions are part of the explanation of the possibility of massive empirical dubitability: The ability to consider and bring into doubt nearly all of our empirical beliefs (though not all at once).

Chapter 1

Fodor's New Radical Conclusion

Introduction

Fodor's original argument against the possibility of concept learning is presented in *The Language of Thought* (1975, hereafter *LOT1*). Since *LOT1* Fodor has revisited and reflected on the implications of the argument in *LOT1* in several places (see Fodor 1981 and 1998 in particular). His stated goal in *LOT2: The Language of thought revisited* is "to summarize, sharpen, and extend those intervening discussions." (2008, hereafter *LOT2*, p. 130) Fodor describes *LOT2* as a "sort of progress report on our continuing effort to inch towards a theory of the intentionality of cognition: what's been done, what still needs to be done, and what, as things stand, we have no idea at all of how to do." (*LOT2*, p.22)

This chapter is a careful attempt at tracing Fodor's path to what I call his "new radical conclusion". That conclusion is the claim that concept acquisition is not a cognitive process, is not in the domain of cognitive psychology, does not admit of psychological explanation, and cannot be explained at the intentional level. The discussion will stay close to its primary text, *LOT2*.

The chapter will proceed as follows: I will begin in section 1.1 with a short primer on Fodor's framework for theorizing about psychological states and process. In section 1.2 I introduce and explain Fodor's "original" and "revised" arguments against the possibility of concept learning. In section 1.3 I introduce Fodor's new radical conclusion in the context of his own positive proposal for a concept acquisition mode. In section 1.4 I propose a general reconstruction of Fodor's argument for the new radical conclusion. I consider four variations of

the general argument and explain how they fit together. Finally in section 1.5 I conclude by reviewing and reflecting on the analysis of Fodor's path to his new radical conclusion.

1.1 Fodor fundamentals

To understand Fodor's arguments we must understand the fundamentals of the Fodorian research program. The foundation of Fodor's research program includes three components: Psychological/intentional realism, naturalism, and the representational/computational theory of mind. I will briefly explain each.

1.1.1 Psychological/intentional realism

Psychological/intentional realism, (which I will refer to simply as intentional realism) acknowledges that intentional states are real and psychological kinds, law-like generalizations, and explanations are couched in intentional vocabulary. On an intentional-realist psychology, behavior is explained in terms of mental causes. For example, if you want to explain why Omar walked into the kitchen and put the kettle on the stove, you'll need to know what Omar desires and what he believes about kitchens and kettles.

1.1.2 Naturalism

Fodor's naturalism is a non-reductive kind of physicalism. On this view, mental states and processes are part of the physical world. This means, at a minimum, that the processes cognitive science postulates "must be ones that can be carried out by actual physical mechanisms, and the states that it postulates are ones that physical objects can be in." (Fodor and Pylyshyn 2014, pp. 4-5)

1.1.3 The representational and computational theory of mind

Squaring intentional realism with naturalism is the driving force of Fodor's research program. In service of that goal Fodor developed the representational and computational theory

of mind. The desired result is a theory of mind that respects the intentionality of propositional attitudes and the causal powers of mental states.

The representational theory of mind (RTM) is a view about the metaphysics of cognitive states and processes. According to RTM “tokens of mental states are tokens of relations between creatures and their mental representations.”¹ (*LOT2*, pp. 5-6) So for example, suppose that at some particular time *t*, Omar believes that the kettle is in the kitchen. According to RTM, at time *t*, there is a corresponding event that consists in Omar being in “some characteristic relation” to a token mental representation that has the semantic content expressed by THE KETTLE IS IN THE KITCHEN.² (1998, p. 8) For simplicity we can think of the characteristic relation as the belief having relation.

Importantly, according to RTM, token mental representations are symbols. And, in turn, tokens of symbols are physical objects. This is RTM’s (and Fodor’s) way of explaining how physical objects can have semantic properties. The computational theory of mind (CTM) is a theory about mental processes. According to CTM, “Tokens of mental processes are ‘computations’; that is, causal chains of (typically inferential) operations on mental representations.” (*LOT2*, pp. 5-6) Demonstrative reasoning is Fodor’s paradigmatic mental process. So, to continue with our example of Omar, suppose that at some time *t*, Omar holds beliefs that express the following content:

(B1) If I want the kettle and the kettle is in the kitchen, then I should go to the kitchen

(B2) I want the kettle and the kettle is in the kitchen

Being of sound mind, Omar infers:

(B3) I should go to the kitchen

¹ “Or, if you prefer, they’re relations between creatures, their mental representations, and propositions that their mental representations express.”

² I use the convention of small caps to mention concepts and complex mental representations such as beliefs. My own view is that concepts are representational contents. However, for simplicity of exposition, given the purposes of this chapter, I adopt RTM’s notion of a concept.

Suppose that B_1 , B_2 , and B_3 are each mental representations that express the corresponding contents above. RTM and CTM explain Omar's kitchen-kettle reasoning process as follows: At time t , Omar stands in the belief-having relation to a token of B_1 and a token of B_2 and the token mental process that takes Omar from having a token of B_1 and a token of B_2 to having a token of B_3 is a token computation, which is a causal relation that respects the semantic properties of the three token mental representations. In this way, because tokens of mental representations are symbols, and tokens of symbols are physical objects, RTM and CTM explain how a physical object such as Omar's neural machinery can token a truth-preserving mental process such as demonstrative reasoning.

Intentional realism, naturalism, and RTM/CTM are the heart of Fodor's research program. But to understand Fodor's arguments in *LOT2*, we will need the basics of Fodor's theory of content. In the intervening years between *LOT1* and *LOT2* Fodor has held various theories of content. For present purposes I will just assume the version of his information-based atomism as it is presented in *LOT2*. Below I outline and explain that theory's commitments most germane to understanding Fodor's discussion of concept acquisition:³

(1) Having a concept is being locked on to the property that the concept expresses.

On Fodor's view, a concept has its content in virtue of getting nomologically locked to a property in the world, thereby coming to carry information about that property.

(2) The extension of a concept is the set of (actual or possible) instances of the property to which the concept is locked. (*LOT2*, p. 141)

(3) Concept possession is atomistic. In principle, one could have some concept C without having any other concept (unless C is a complex concept in which case one must have C 's constituent concepts). (*LOT2*, p. 141)

³ The three commitments I discuss are drawn from a larger list. See, in particular, *LOT2*, p. 141.

So for example, on Fodor's view, acquiring the concept CAT is getting locked on to the property of being a cat (P_{cat}) such that tokens of P_{cat} cause tokens CAT. The extension of CAT is the set of actual or possible instances of P_{cat} . And one can, in principle, have the concept CAT without having any other concept.

1.2 The Arguments of *LOT1* and *LOT2*

I'll begin by presenting Fodor's arguments in *LOT1* and *LOT2*—what I'll call “the original argument” and “the revised argument,” respectively. I will discuss the reasons Fodor marshals in favor of each premise, and then I will note the differences between the revised argument and Fodor's original argument for radical concept nativism in *LOT1*. I'll call the argument in *LOT1* “the original argument”.

The Original Argument (*LOT1*, Ch.2)

(OA1) All concepts are either learned or innate.

(OA2) All concepts are either a definitional primitive or constructed out of definitional primitives.

(OA3) If a concept is acquired by learning, then it is acquired by the process of hypothesis formation and confirmation.

(OA4) If a concept is a definitional primitive, then it can't be acquired by hypothesis formation and confirmation.

(OA5) So if a concept is a definitional primitive, then it can't be learned.

(OA Conclusion) So if a concept is a definitional primitive, then it must be innate.

What makes the conclusion radical is that it turns out that the number of concepts that are definitional primitives—often referred to as “lexical concepts”—is quite large. For example, by one estimate there are 500,000 entries in the *Oxford English Dictionary* that correspond to lexical concepts. Lexical concepts are monomorphemic concepts—concepts that correspond to morphemes, typically characterized as the smallest meaningful units of a natural language. Monomorphemic concepts contrast with complex concepts. Complex concepts are concepts that decompose into monomorphemic concepts. For example the concept GREEN SHOE decomposes into two monomorphemic concepts GREEN and SHOE. The conclusion of Fodor's original

argument won him the title of radical concept nativist. What makes the conclusion so radical is not only the number of concepts that it posits as innate, but also the kinds of concepts that would be included in the innate inventory. Famous examples include QUARK and CARBURETOR—concepts that are beyond the pale of what we could plausibly attribute to our ancient ancestors.

Let us turn now to the revised argument of *LOT2*:

The Revised Argument (*LOT2*, Ch.5; *SSFPCA*, 2005)

(RA 1) If there is such a thing as concept learning, concept learning is a process of hypothesis formation and confirmation.

(RA 2) (On pain of circularity) Concept learning cannot be a process of hypothesis formation of hypothesis formation and confirmation.

(The Revised Argument Conclusion) There is no such thing as concept learning.

The first premise of the revised argument (RA1) corresponds to the third premise in the original argument (OA3). The premises correspond to what Fodor describes as the following “common ground principle”:

“*Concept learning* is a process of inductive inference; in particular ...it’s a process of projecting and confirming hypotheses about what the things that the concept applies to have in common.” (*LOT2*, p. 131)

Fodor calls this the “HF model” of concept learning. Fodor claims the consensus over the common ground principle is general but rarely acknowledged (*LOT2*, p. 132).

He speculates that, “this unhappy state of affairs results from a pretty general failure to recognize that if concept acquisition is a learning process then (assuming RTM) it requires the mental representation of the conditions under which the concept applies (as in: ‘It’s the green things that GREEN applies to’).” (*LOT2*, p. 133)

Fodor points to RTM as the source of the common ground principle. According to Fodor, being committed to RTM commits one to the claim that learning a concept is a species of belief acquisition. Thus acquiring the concept GREEN by a learning process requires acquiring the belief that:

(B_{GREEN}) It's the green things that GREEN applies to.

Fodor insists that (assuming concept learning is possible) the fact that it is possible to learn one but not the other of “two distinct but coextensive concepts” shows that concept learning is some kind of hypothesis formation and confirmation (*LOT2*, p. 135; *LOT1*, pp. 95-96). He writes,

“The learning processes that distinguish between coextensive concepts require representing the same experiences in different ways. (This is a consequence of assuming RTM is true of concept learning; that is that experience affects concept learning only as it is mentally represented.)” (*LOT2*, p. 135)

However, it is not clear that his reasoning turns on the fact that it is possible to learn one but not the other of two distinct but coextensive concepts. His reasoning is as follows:

(HF1) If you can acquire one of two distinct coextensive concepts C and C^* on the basis of experiencing an x that is an instance of C and C^* , then there must be at least two ways of representing x on the basis of which C or C^* can be learned.

(HF2) You can acquire one of two distinct coextensive concepts such as C and C^* on the basis of experiencing an x that is an instance of C and C^* .

(HF3) There must be at least two ways of representing x on the basis of which C or C^* can be learned.

(HF4) If concept acquisition is a learning process, then (assuming RTM) it requires the mental representation of the conditions under which the concept applies. E.g., learning the concept C requires learning that it's the C things that C applies to.

(HF5) Hypothesis formation and confirmation is the only mental process by which a subject proceeds from representing some things as C to representing all such things as C s (*LOT2*, 134).

(HF-Conclusion) If concept acquisition is a learning process, then it must be a process of hypothesis formation and confirmation.

If this is indeed Fodor's reasoning then it's unclear what work the possibility of learning one of two coextensive concepts is actually doing. That is, if you grant HF4 and HF5, then it follows that if there is such a thing as concept learning, it must be a process of hypothesis formation and confirmation.

So it appears that on Fodor's view (which assumes RTM), acquiring a concept is getting locked on to the property that a concept expresses, but learning a concept is acquiring a belief about what the concept applies to. Unsurprisingly, a number of philosophers and psychologists

have doubted that concept learning must proceed by hypothesis formation and confirmation.⁴ In response, Fodor diagnoses an epidemic in cognitive psychology whereby theorists “accept HF without fully realizing that it’s HF that they accept.”⁵ (*LOT2*, p. 132) What Fodor has failed to adequately defend (let alone explain) is why HF is a consequence of assuming RTM. RTM is an exceedingly more plausible candidate for being a “common ground principle.” Barring a convincing argument for that claim that RTM requires that concept learning must be a process of hypothesis formation and confirmation, or some argument independent of RTM, Fodor’s claim that HF is common ground is at best mysterious, and at worst simply false.

Let us turn to the second premise in the revised argument. Recall that RA2 of the revised argument is:

(RA2) (On pain of circularity) Concept learning cannot be a process of hypothesis formation and confirmation.

Fodor thinks that if there were such a thing as concept learning then it would be the processes of forming and confirming hypotheses. What would such a hypothesis look like? The following is an example of the form that Fodor has in mind:

(HYP_{GREEN}) It’s the green things that GREEN applies to. (*LOT2*, p. 133)

Notice that the hypothesis not only mentions the concept GREEN but also makes use of the concept GREEN. So the hypothesis presupposes the possession of the concept GREEN. If the

⁴ For various criticisms of Fodor’s restricted use of the term ‘learning’ see Samet and Flanagan (1989), and Margolis and Laurence (2011). For some discussion on the heterogeneous nature of the use of the term ‘learning’ in cognitive science and psychology see (for example) Spelke and Newport (1998), Keil (2008), and Gallistel and King (2009). For replies by psychologists specifically directed at Fodor’s notion of learning see Carey (2009) and (SSFPCA, 2005). For criticisms that specifically address Fodor’s attempts at drawing the distinction between learning processes and triggering or merely brute-causal processes see Sterelny (1989), Samet and Flanagan (1989), and Weiskopf (2008).

⁵ Fodor writes, “Over the years, a truly remarkable number of cognitive scientists who think that concept learning is a kind of inductive inference have assured me that they do not think that concept learning is a kind of inductive inference. (As I write this, I’m just back from a large cognitive science conference where almost everybody thought that concept learning is a kind of inductive inference and almost everybody denied that they did. My nerves are in terrible shape.)” (*LOT2*, p. 132)

hypothesis presupposes the possession of the concept GREEN then it cannot explain the acquisition of GREEN. So concept acquisition by hypothesis formation and confirmation is not possible. This reasoning is rehashed several times in *LOT2*'s chapter on nativism. Here is just one of many examples:

“Quite generally, you can't represent anything as such and such unless you already have the concept such and such. All that being so, it follows, on pain of circularity, that 'concept learning' as HF understands it can't be a way of acquiring concept C. For, the one thing that a process whereby C is acquired can't be is a process that presupposes having C. Conclusion: If concept learning is as HF understands it, there can be no such thing.” (*LOT2*, p. 139)

There are two noteworthy differences between the original argument of *LOT1* and the revised argument of *LOT2*. The original argument treated the notions of being innate and being learned as exhaustive. Upon reflection, Fodor admits that it doesn't follow from the argument in *LOT1* that any concepts are innate because concepts can be acquired by processes other than learning (*LOT2*, p. 130). Fodor clarifies that a concept acquisition process is “any process that adds to the conceptual repertoire”, and so to claim that a concept is innate is to claim that it is not acquired. (*LOT2*, p.132)

The revised argument differs from the first in another important way. The argument in *LOT1* applied only to lexical concepts. In his original argument Fodor's target was the view that learning a concept is learning its definition. That is why his argument focused on the fact that lexical concepts are definitional primitives. Complex concepts were spared the consequence of the original argument because Fodor thought that complex concepts could be learned through the logical construction of two or more definitionally primitive concepts (*LOT1*, p. 96).

In *LOT2* Fodor notes that the claim that concept learning—if possible—must be circular, applies to all concepts, monomorphemic and complex. This revision is correct with respect to Fodor's own argument for the circularity of concept learning. The problem that Fodor poses for

monomorphemic concepts holds just as well for complex concepts. Recall that the form of a “concept learning” hypothesis is the following:

(HYP_{GREEN}) It’s the green things that GREEN applies to.

Compare this with the hypothesis for the complex concept green shoe:

(HYP_{GREEN SHOE}) It’s the green shoe things that GREEN SHOE applies to.

The hypothesis for learning complex concepts also presupposes the concept that is to be learned.

So the argument against concept learning is general: It applies to all concepts, monomorphemic and complex. Fodor characterizes this as a stronger conclusion and submits that, “What I should have said is that it’s true and a priori that the whole notion of concept learning is *per se* confused.

Punkt!” (LOT2, p. 130)

1.3 Fodor’s new radical conclusion

While memorable, the slogan “the whole notion of concept learning is *per se* confused. *Punkt!*” is not the shocking headline of Fodor’s progress report in LOT2. The real headline is the much stronger thesis that concept acquisition is not in the domain of cognitive psychology and does not admit of psychological explanation. I’ll call this thesis “Fodor’s new radical conclusion”. What is at issue is not merely a terminological dispute. Fodor claims that concept acquisition is a process that does not admit of psychological explanation and can only be given a neurological explanation. Understanding how Fodor reaches his new radical conclusion will take some exegetical work. The following passage shows what he takes to be the central issue given his revision of the argument in LOT1. He writes,

“I do think that if HF is no longer in the running, then a general reformulation of the issues about concept innateness is mandatory. The central issue isn’t which concepts are learned, since, if the (emended) LOT1 argument is right, none of them are. Nor, however, is it which concepts are innate, if an innate concept is one the acquisition of which is independent of experience. Quite likely, there are none of those either. Rather, the problem is to explain how a creature’s innate endowment (whether it is described in neurological or in intentional terms) contributes to the

acquisition of its conceptual repertoire; that is, how innate endowments contribute to the processes that start with experience and end in concept possession.” (*LOT2*, p. 145)

I propose that we accept Fodor’s characterization of the explanatory task at hand. The rest of this chapter will be dedicated to trying to understand why it is that Fodor decides that the explanation cannot be given in intentional terms. In the service of that goal I’ll begin by describing Fodor’s sketch of a possible concept acquisition process that starts from experience and ends in concept possession.

1.3.1 Fodor’s positive proposal

On Fodor’s positive proposal, the concept acquisition process begins with the forming of a statistical representation that he calls a “stereotype.” Fodor does almost nothing by way of explicating the notion of a stereotype that he employs. He says that stereotypes are “statistical representations of experience,” (*LOT2*, p. 162) that “stereotypic instances of a concept are the ones that most resemble most other instances of the concept” (*LOT2*, p. 149) and that stereotypic properties are “central tendencies of (actual or possible) populations.” (Thus Fodor must be assuming that the concepts that make up stereotypes are already present).

Fodor summarizes his positive proposal using the following diagram:

“Initial state \rightarrow (P1) \rightarrow stereotype formation \rightarrow (P2) locking (= concept attainment).” (*LOT2*, p. 151)

According to Fodor, P1 is a process of statistical inference by which a mind learns a stereotype. P2 is the process “that transforms a mind that has learned a stereotype into a mind that has acquired the concept whose stereotype it is.” (*LOT2*, p. 154) It is the second process that Fodor claims is out of the domain of cognitive psychology and cannot be given a psychological explanation. He writes, “P2 is some reliable but not intentional (and hence, a fortiori, not inferential) neurological process.” (*LOT2*, p. 151) So while stereotypes are learned by induction, concepts are acquired from stereotypes by some neurological process.

1.3.2 A closer look at Fodorian stereotypes

Before looking further into Fodor's positive proposal, I should address a pressing question about stereotypes and stereotype learning. Namely: What are stereotypes such that they can be learned given Fodor's view of concept learning? As I indicated above, Fodor says almost nothing about what he takes stereotypes to be, but claims that stereotypes can be learned by induction.

What follows is my attempt at giving the most charitable and plausible guess at what Fodor has in mind. Fodor's stereotypes are clearly complex representations, but we should resist the temptation to think of them as complex concepts. Rather, stereotypes are more like beliefs that carry statistical information about the co-instantiation of properties in the world. Specifically, they are beliefs about which properties or features tend to go together. On the most plausible version of this interpretation the relevant properties would be superficial properties. For example, consider what could be the cat stereotype. Suppose that the stereotypic features of cats include some characteristic shape, coloring, texture, motion pattern, and sound. Call these features 1-5. For simplicity I will call the characteristic shape of cats the "x-shape" and the characteristic motion pattern of cats "the x-biological-motion." Furthermore, let us suppose that the characteristic coloring of cats is greyish-brown, that the characteristic texture of cats is furry, and that the characteristic sound made by cats is a meowing sound. On the interpretation of Fodor's view that I am proposing the concepts that indicate these characteristic features are available for the construction of hypotheses.⁶ These concepts could be innate or could be acquired through maturation (or some other process) but not by anything that Fodor would call a

⁶ Fodor is a conceptualist about perceptual attributives so I am assuming that the constituents of stereotypes are conceptual.

“learning process.” Thus on the interpretation I am offering the hypothesis that must be confirmed in order to learn a stereotype would be something like:

(HYP-stereotype_x) Being x-shaped, greyish-brown, furry, moving in the x-biological motion way, and making meowing sounds, tend to go together.

It may be that you have not learned a stereotype until you have reached a certain level of confirmation or what is known as a “criteria for learning” in psychology experiments. Thus in learning a stereotype you do not acquire a *new* way of representing things. What you have done is confirmed a hypothesis about the co-instantiation of properties. On this interpretation learning a stereotype is more like solving a sorting problem than acquiring a concept. The sorting that needs to be figured out is which properties tend to go together. On this interpretation stereotypes are learned by induction because the confirmation of hypotheses is an inductive process.

To be clear, I am not advocating for this notion of a stereotype or for Fodor’s positive proposal; and I am certainly not endorsing Fodor’s claims about learning. I offer the above interpretation of Fodorian stereotypes as an attempt to give the most charitable interpretation of Fodor’s positive proposal. To that end I have tried to construct an explanation of stereotype learning that is at least consistent with Fodor’s own arguments against the possibility of concept learning.⁷

⁷ Eric Margolis and Stephen Laurence (2011) give an alternative to the interpretation I’m giving here. On their interpretation stereotypes are analogous to complex concepts. Thus they argue that Fodor’s own view suggests that learning a concept’s stereotype would require putting forward a hypothesis about the stereotype’s individuating conditions, and that this makes Fodor subject to his own argument against the possibility of concept learning. I agree that this is the most natural reading of the little that Fodor says about learning a stereotype. Given their reading, Margolis and Laurence confront Fodor does with his own self-made dilemma: Either stereotypes cannot be learned because of Fodor’s own circularity argument, or there is some way other than hypothesis testing by which one can learn a stereotype. If the latter is the case then Fodor must explain why concepts cannot be learned in this “other way.” (2011, p. 533)

1.3.3 A closer look at P2: Acquiring a concept as a consequence of learning a stereotype

Fodor considers the question: how does one get from (a representation of) the extension of the stereotype to (a representation of) the extension of the corresponding concept? He answers:

“It’s no sort of inference at all. It’s... a *subintentional* and *subcomputational* process; it’s a kind of thing that our sort of brain tissue just does. Psychology gets you from the initial state to P2; then neurology takes over and gets you the rest of the way to concept attainment (that is, to locking according to the present assumptions).” (*LOT2*, pp. 151-152)

So according to Fodor’s positive proposal stereotypes are learned, concepts are “merely acquired,” and what is innate is “The disposition to grasp such and such a concept (i.e. to lock to such and such a property) in consequence of having learned such and such stereotype”.⁸ (*LOT2*, p. 162) A creature having a disposition to grasp a concept C as a consequence of learning a stereotype *s* “is a brute fact about that creature, (or perhaps that kind of creature).” (*LOT2*, p. 163)

Fodor concludes Chapter 5 of *LOT2* by writing “the pressing research problem now is to understand the relationship between concepts and stereotypes”... “why is it that learning the C stereotype triggers the acquisition of C and not some other concept (equivalently why it triggers the locking to the property that C expresses and not some other property). (*LOT2*, p. 168) Fodor suggests that processes such as imprinting and parameter setting may be analogous processes in the following sense: they are processes that are “highly sensitive to the character of experience”, but it is implausible that they are achieved “by rational processes like inductive inference.”⁹ (*LOT2*, pp. 144-5)

⁸ On this story it turns out that the difference between radical and modest concept nativism lies “in the generality of the principles that map the stereotypes onto the concepts.” (*LOT2*, p. 162)

⁹ Of course it should be noted that imprinting and parameter setting are notions that were introduced and studied by psychologists. If nothing else, this suggests that psychology has at least some role to play in the explanation of such processes. So even if the analogy holds it does not support Fodor’s new radical conclusion.

Thus Fodor ultimately concludes that P2 cannot “be explained at the intentional level”. (*LOT2*, p. 163) On Fodor’s view there is no explanation above the level of neurology of how stereotypes lead to the acquisition of concepts. If the only explanatory account of the relationship between a stereotype and the concept acquired on the basis of that stereotype is located at the neurological level, then the law-like generalizations that supposedly capture the regularities of such transitions will be couched entirely in the non-intentional vocabulary of neuroscience.

1.4 Fodor’s path to the new radical conclusion

The rest of the chapter will be dedicated to trying to figure out how Fodor arrives at the new radical conclusion. I will begin by clarifying the scope of the new radical conclusion. Based on what he says in his chapter on nativism (*LOT2*, Ch. 5) I read Fodor as accepting the following:

If m is a cognitive process, then m is in the domain of cognitive psychology, m admits of a psychological explanation, and an explanation of m is an explanation at the intentional level.

And what I have been calling “Fodor’s new radical conclusion” is the claim that:

Concept acquisition is not a cognitive process, is not in the domain of cognitive psychology, does not admit of psychological explanation, and cannot be explained at the intentional level.

1.4.1 Fodor’s argument in its most general form

In fact what I have been calling the “new radical conclusion” has been lingering in Fodor’s work since *LOT1* (1975). Since its publication four decades ago, Fodor has written extensively about concepts, cognitive science, and concept acquisition. (See especially 1980, 1981, 1998, 2000, and 2008 (*LOT2*)). There are several arguments that can be extracted from these writings. The arguments frequently run into and through each other. In the service of dialectical housekeeping I will reconstruct Fodor’s argument for the new radical conclusion in its most general form as an argument *schema*. I call it “Fodor’s New Radical Conclusion Argument Schema” (or simply NRC):

NRC Argument Schema

(1) If concept acquisition is a cognitive process, is in the domain of cognitive psychology, admits of psychological explanation, and can be explained at the intentional level, then concept acquisition has property p*.

(2) It is not the case that concept acquisition has property p*.

(3) So concept acquisition is not a cognitive process, is not in the domain of cognitive psychology, does not admit of psychological explanation, and cannot be explained at the intentional level.

We can distinguish between the various arguments Fodor invokes by specifying some property p* and substituting that specification in place of “has property p*” in premises (1) and (2) of the NRC schema above. Fodor’s writings suggest at least 6 possible specifications of p*. They are the following:

(a) is a process of hypothesis formation and confirmation.

(b) is a learning process.

(c) is a computational process.

(d) is a process where the relevant acquisition transition is a transition between mental states that preserves truth and/or semantic content.

(e) is a process where some experience is treated as evidential for the acquired concept.

(f) is an inferential/rational process.

Using (a)-(f) and the NRC argument schema we can produce six different versions of the argument for the new radical conclusion. I will refer to the arguments as NRC(a), NRC(b), NRC(c), etc. For ease of exposition I will shorten “is a cognitive process, is in the domain of cognitive psychology, admits of psychological explanation and can be explained at the intentional level” to “is a cognitive process (...).”

NRC(a) and NRC(b) are closest to Fodor’s revised argument against concept learning.

Recall that I formulated the revised argument as follows:

The revised argument

(RA 1) If there is such a thing as concept learning, concept learning is a process of hypothesis formation and confirmation.

(RA 2) (On pain of circularity) Concept learning cannot be a process of hypothesis formation and confirmation.

(The Revised Argument Conclusion) There is no such thing as concept learning.

Compare the revised argument to NRC(a) and NRC(b):

NRC(a)

(a1) If concept acquisition is a cognitive process (...), then concept acquisition is a process of hypothesis formation and confirmation.

(a2) It is not the case that concept acquisition is a process of hypothesis formation and confirmation.

(a3) So concept acquisition is not a cognitive process (...).

NRC(b)

(b1) If concept acquisition is a cognitive process (...), then concept acquisition is a learning process.

(b2) It is not the case that concept acquisition is a learning process.

(b3) So concept acquisition is not a cognitive process (...).

Because much of the dialectic surrounding the original and revised arguments has focused on Fodor's characterization of learning as hypothesis formation and confirmation, I will concentrate on the four other ways of specifying p^* in the argument schema for the new radical conclusion:

(c) is a computational process

(d) is a process where the relevant acquisition transition is a transition between mental states that preserves truth and/or semantic content.

(e) is a process where experience is treated as evidential for the acquired concept.

(f) is a rational/inferential process

In what follows we will see that versions NRC(c)-(f) of the argument for the new radical conclusion rely on arguments from the underdetermination of concepts by stereotypes.

1.4.2 The projection problem and NRC(c)-(f)

I propose that we proceed by considering the arguments for the new radical conclusion in the case where the process in question is P2, the process that takes a mind from learning a stereotype to concept attainment. What is at issue then is whether P2 has properties (c)-(f).

Fodor claims that P2 cannot be an inferential process because there are "so many different kinds of relations in which a concept and its stereotype may stand." (*LOT2*, p. 154)

Fodor points to the underdetermination of concepts by stereotypes saying that we cannot "just look at a stereotype and figure out what concept it is the stereotype of..." He adds that, "the

underdetermination of concepts by their stereotypes looks to be very radical.” (*LOT2*, p.155) I call the underdetermination of concepts by stereotypes “the projection problem.”¹⁰

Fodor concedes that a stereotype may very well be learned, and that the learning of a stereotype can play a causal role in the acquisition of a concept. However, Fodor rejects the possibility that the transition from a stereotype to a concept is some kind of psychological transition because, as he says, “there are so many different kinds of relations in which a concept and its stereotypes may stand” (*LOT2*, p. 154). Because stereotypes do not wear a unique concept on their semantic sleeve, there must be some account of how some stereotype *s* leads to the acquisition of a particular concept *C*. Fodor’s claim is that there could not be a single account that characterizes which stereotypes lead to the acquisition of which concepts, because there is no single relation that illuminates why some stereotype *s* leads to the acquisition of some concept *C* as opposed to some other concept *C**. Specifically, he explains, there is no “pattern of inference which, given just a stereotype, identifies the corresponding concept.” (*LOT2*, p. 154)

Consider, for example, the underdetermination of the concept *CAT* by the stereotypic properties of cats. For simplicity, let us again use the stereotypic features of cats that I called features 1-5 above. The cat stereotype underdetermines the concept *CAT*, because one could project to a multitude of different concepts based on the features included in the cat stereotype. Obviously features 1-5 are stereotypical of the extension of *CAT* *and* the entities in the extension of the concept *ANY THING THAT EXHIBITS FEATURES 1-5*. Call this latter concept the *SHMAT* concept.¹¹ Although they have different extensions, *CAT* and *SHMAT* can share the same

¹⁰ In the present chapter I focus on Fodor’s use of the projection problem to argue for the new radical conclusion. In Chapter 2 I explain how Eric Margolis’s (1998) concept acquisition model resolves the projection problem for natural kind concepts.

¹¹ Recall that on the charitable reading of Fodor’s positive proposal stereotypes are hypotheses that have reached some level of confirmation. We can think of these as beliefs that have reached some credence criterion. So (once again) for example, what I am calling the “cat stereotype” is the belief that: Being x-shaped, greyish-brown, furry,

stereotype. Abnormally looking and abnormally moving cats are included in the extension of CAT. However, these unfortunate souls may be too deviant to make it into the extension of SHMAT. Moreover, cat-fakes—(non-cats that exhibit features 1-5)—will be in the extension of SHMAT. Putnam’s famous scenario in which all of the entities on Earth that we think of as cats actually turn out to be robots controlled by radio from Mars (hereafter simply “robocats”) provides yet another example. The concepts ROBOCAT and CAT have disjoint extensions but share the same stereotype. (Putnam, 1962)

The preceding examples show that concepts with different extensions can share a stereotype. But the projection problem also arises in cases of coextensive but intentionally distinct concepts. To use a variation of one of Fodor’s favorite examples, suppose cats are your grandmother’s favorite animal. The cat stereotype will also be the stereotype produced by instances of the property of being Granny’s favorite animal. Thus to overcome underdetermination, we need an explanation of why the cat stereotype leads to the acquisition of CAT rather than the acquisition of GRANNY’S FAVORITE ANIMAL. Moreover, given what is included in a stereotype, why should the cat stereotype lead a child to acquire the concept CAT as opposed to the concept ROBOCAT or the concept GRANNY’S FAVORITE ANIMAL? Another way to think of the problem poses the following question: What general pattern of inference, given just a stereotype, identifies the corresponding concept. What is at issue for Fodor is specifying the dispositions operative in concept acquisition that could, in combination with learning a stereotype, lead to the acquisition of a concept.

moving in the x-biological motion way, and making meowing sounds, tend to go together. Having this belief is different from having the concept THINGS THAT EXHIBIT FEATURES 1-5 (aka the concept SHMAT). However, it may be the case that having the former always leads the acquisition of the latter. Whereas having the concept SHMAT need not automatically lead to having the belief that being x-shaped, greyish-brown, furry, moving in the x-biological motion way, and making meowing sounds, tend to go together.

Fodor invokes the underdetermination of concepts by stereotypes to argue that the disposition to acquire some concept *C* in consequence of having learned the *C* stereotype is just a brute neurological fact. Thus the constraints that shape the dispositions are not specifiable or explainable at the intentional level. Fodor concludes that we are just hardwired to move from a stereotype to the concept that it is a stereotype of. So, for example, we are just hardwired to move from the *CAT* stereotype to the acquisition of the concept *CAT* as opposed to some other concept. It is, he writes, “a kind of thing that our brain tissue just does.” (*LOT2*, p. 152)

Fodor equivocates between different kinds of underdetermination. Specifically, his comments suggest that the projection problem exemplifies both a semantic underdetermination and an epistemic underdetermination. In both cases Fodor invokes underdetermination to try to show that *P2* is neither rational nor inferential. I discuss epistemic and semantic underdetermination in the subsequent two sections. As we shall see, what will begin to emerge is a larger and even more controversial radical thesis whereby many inductive processes, specifically abductive processes, suffer the same fate as the process of concept acquisition.

1.4.3 Epistemic underdetermination, the argument from arbitrariness, and Fodor’s brute/rational dichotomy

Fodor’s use of epistemic underdetermination turns on his assumption that the experience from which stereotypes are learned would have to be evidential. He treats being evidential as the missing epistemic relation between experience and the acquisition of concepts. Fodor’s thinks the purported evidential relation between experience and concept acquisition is why concept acquisition has the look of a rational process. Consider the following passages:

[T]he experience from which a concept is learned must provide (inductive) evidence about what the concept applies to. Perhaps *COW* is learned from experience with cows? If so, then experiences with cows must somehow witness that it’s cows that *COW* applies to. This internal connection between concept learning and epistemic notions like *evidence* is the source of the strong intuition that concept learning is some sort of rational process. It contrasts sharply with

kinds of concept acquisition where, for example, a concept is acquired by surgical implantation or by “swallowing a pill; or by hitting one’s head against a hard surface, etc. Intuitively, none of these is concept learning; but any or all of them might eventuate in concept attainment. (*LOT2*, pp. 135-136)

And,

“If the relation between concepts and experiences is typically evidential, that would explain why it’s so often a relation of content: and what other explanation have we got?” (1998, p. 132).

In his early work criticizing the possibility of concept learning Fodor writes,

“...the distinguishing characteristic of concept *learning*, in the *non-technical, pre-theoretic* sense of that notion, is the existence of a rational relation between the concept acquired and the experience that occasions its availability...” (1981, p. 275)

Fodor contrasts the “rationality” of concept learning with merely “brute-causal” processes whereby an experience causes the acquisition of a concept. In the case of merely brute-causal processes, the relations between an experience and the concept acquired “may often be *extremely arbitrary*” (1981, p. 280).

Fodor frequently slides from lacking evidence for acquiring a concept to the arbitrary and thus irrational and non-inferential nature of the acquisition process. Moreover, Fodor moves from the arbitrary and irrational nature of the concept acquisition process to the claim that the transition from experience to concept acquisition is merely “brute-causal.” Thus Fodor seems to leave out the quite plausible possibility that concept acquisition is neither brute-causal nor rational. Finally, in what may be the least controversial of the moves he makes, Fodor moves from the claim that concept acquisition is a brute-causal process to the new radical conclusion: Concept acquisition is not a cognitive process, is not in the domain of cognitive psychology, does not admit of psychological explanation, and cannot be explained at the intentional level.

1.4.4 Semantic underdetermination and computational processes

I said that Fodor equivocates between epistemic underdetermination and semantic underdetermination. Understanding his argument from semantic underdetermination to the new radical conclusion is a complicated task. I will begin by discussing three of the six properties that provide a way of specifying the argument for the new radical conclusion in tandem. They are the following:

(c) is a computational process.

(d) is a process where the relevant acquisition transition is a transition between mental states that preserves truth and/or semantic content.

(f) is an inferential/rational process.

Fodor considers the question “Who says concept learning is an *inferential* process at all? Perhaps it’s just some kind of brain tickle.” He replies,

“[I]’m assuming at this point that concept learning is in the domain of cognitive psychology, hence (according to CTM) that it’s some sort of computational process; a process in which computations take a mind from premises to conclusions preserving such semantic properties as content and/or truth.” (*LOT2*, p. 137)

Recall that according to CTM, a computation is a causal relation between symbols that respects semantic content. The goal of CTM is to explain how rational processes can be mechanically implemented such that mechanical processes can respect truth-preserving inferences among mental states. As I noted above, a demonstrative inference such as *modus ponens* is the canonical inference pattern that Fodor trots out when singing the praises of CTM.

Let us consider whether P2 is a process that can be said to preserve semantic content and/or truth. I will consider the preservation of semantic content first.

In some of Fodor’s more detailed discussions of computational processes and the preservation of semantic content, Fodor seems to be largely concerned with a specific content preservation problem: The substitution of co-extensive concepts in opaque contexts such as

propositional attitudes. The canonical cases are Frege cases that involve co-referential proper names such as the case of ‘Wonder Woman’ and ‘Diana Prince’. A substitution of the concept WONDER WOMAN for the concept DIANA PRINCE in an opaque context does not preserve semantic content because WONDER WOMAN and DIANA PRINCE differ in their semantic content.¹² Thus the beliefs (B4) Wonder Woman has super strength, and (B5) Diana Prince has super strength, differ in semantic content even though they attribute the same property to the same individual.

Fodor’s discussions of the problem that Frege cases pose for computational processes give us some context for his claim that computational processes must preserve semantic content. However, it is difficult to link those discussions with his meditations on concept acquisition and the impossibility of concept learning.¹³

On the interpretation of Fodorian stereotypes described earlier, the semantic content of stereotypes underdetermines and obviously differs from the semantic content of the concepts to which they project. We do not need to invoke the existence of co-extensive concepts to show this. That the process of concept acquisition cannot be said to preserve semantic content is not surprising. It is why the concept acquisition “puzzle” has so frequently been posed as a puzzle

¹² Fodor (at least since 1994) is a referentialist, so he holds that semantic content is determined purely by reference. Being a referentialist and a holder of RTM/CTM, Fodor’s Frege problems are pushed down one level to the problem of distinguishing the causal powers of co-extensive concepts. His solution is to distinguish co-extensive concepts at the syntactic level, thereby resolving the problem of distinguishing the causal powers of co-extensive concepts through the difference in syntactic properties. This “resolves” Frege problems by the lights of a referentialist. Fodor subscribes to referentialism because he thinks it’s the only option for a committed naturalist. There are avenues by which non-referentialists can push back on his purported “solution.” (See, for example, criticisms by various philosophers in Loewer and Rey, 1991). However, since this chapter is focused on understanding how Fodor reaches the new radical conclusion a discussion of this dialectic would require a considerable digression.

¹³ The one notable exception is Fodor’s argument for the first premise (RA 1) in what I called “the Revised Argument.” (RA 1) reads: If there is such a thing as concept learning, concept learning is a process of hypothesis formation and confirmation. Recall that part of Fodor’s defense included the possibility of learning one of two co-extensive concepts based on an experience of any thing that falls in their extension. But recall also that upon closer inspection we saw that the possibility of learning one of two co-extensive concepts was superfluous to Fodor’s argument.

about the possibility of increases in expressive power. The expressive power of some mind is characterized in terms of all of the representational content that the mind could entertain given the mind's stock of representational resources. Such resources include concepts and the logical and grammatical machinery needed to construct complex representations.¹⁴ Adding a concept increases the mind's expressive power because it adds a way of representing that was previously unavailable. So, for example, if S acquires the concept KOOKABURRA, then the full power of S's representational abilities after acquiring KOOKABURRA cannot be described using only the expressive power of her representational abilities before she acquires KOOKABURRA. To take issue with the semantic underdetermination side of the projection problem is to take issue with the possibility of increases in expressive power.

Fodor's view that you cannot achieve increases in expressive power through psychological processes goes back at least as early as *LOTI*, and his discussions of Jean Piaget's theory of discontinuous developmental stages. Fodor denied and continues to deny "that one can learn a language whose predicates express extensions not expressible by predicates of the representational system whose employment mediates the learning." (*LOTI*, p. 87) Notice that Fodor is not denying the possibility of increases in expressive power full stop. Fodor leaves it open that increases in expressive power could be a result of innately specified maturational processes or a well executed blunt trauma. What he denies is the possibility that an increase in expressive power—such as the increase achieved by the acquisition of some concept C—can be mediated by the representational resources of a mind before it acquires that concept C. That reasoning should sound familiar. It is the same line of reasoning that lead Fodor to declare that there could be no such thing as concept learning.

¹⁴ Again, because Fodor is a conceptualist about perceptual content I will not bother distinguishing between perceptual attributives and conceptual attributives.

Let us turn now to consider whether P2 is a process that can be said to preserve truth. It is difficult to conjure up a construal of the transition from a stereotype to the acquisition of a concept that guarantees the preservation of truth. For example, if the transition is read as embodying the conditional “if being x-shaped, greyish-brown, furry, moving in the x-biological motion way, and making meowing sounds, tend to go together, then there are cats,” it would be false, because it could be the case that our world is overrun by creatures that support the antecedent, but sadly lacks any cats. If that is the construal of the transition that Fodor has in mind, then it starts to seem like Fodor just has a problem with inductive—or at least abductive—inferences. That was the impression suggested by Fodor’s invocation of epistemic underdetermination as well. That Fodor has an underlying problem with accounting for abductive inferences is explicitly discussed in his own reflections on the explanatory challenges facing CTM. (2000, *LOT2*, Ch. 4)

1.4.5 Fodor on computational processes more generally

Thus far we have seen that P2 does not preserve semantic content or truth, and that there is a larger problem for CTM looming in the background. Fodor concedes that CTM is not well equipped to handle non-demonstrative inference such as abduction, and that many of the processes that are of interest to cognitive psychology are of this kind, including perception and learning. (*LOT2*, Ch. 4) However, Fodor’s concession stops short of abandoning CTM because of its marked success explaining demonstrative inferences, and because, as he frequently insists, there are no plausible alternatives. (See especially, 2000 and *LOT2*, Ch. 4).

The problem originates with what Fodor calls the inherent “locality” of computational processes. According to CTM, if mental processes are computations then they are defined over the constituent structure of mental representations. Thus Fodor explains that according to CTM, “computational processes are local” in the sense that “whether a certain mental representation is

in the domain of a certain mental process depends solely on the relation between that representation and its parts” (*LOT2*, p. 108). Demonstrative inferences are local because all of the relevant information needed to draw the inference is present in the internal semantics and syntax of the mental representations that enter into the inference. For example, consider the following simple example of a demonstrative inference:

- (1) If x is a brown cow, then x is brown and x is a cow.
- (2) x is a brown cow.
- (3) So x is brown and x is a cow.

Given the semantics and syntax of (1) and (2), (3) can be inferred without any further information than what is already available in (1) and (2). The problem with non-demonstrative inferences according to Fodor is that they are inherently non-local, that is, (roughly) the internal semantics/syntax of the representations in such processes do not contain all of the potentially relevant information for securing the inference.

According to Fodor’s analysis, context sensitive processes are non-local by definition. (2000, pp. 29-31)¹⁵ Potentially all of one’s cognitive commitments are relevant for such inferences. Any inference to the best explanation is supposed to be an example of a context sensitive inference, because what the “best explanation” is can change from context to context.¹⁶

Giving a thorough exposition of Fodor’s notion of locality is outside the scope of this short summary. But a common thread in the cases of non-demonstrative inferences that concern Fodor is that non-local inferences are locally underdetermined. The transition from a stereotype to the acquisition of a concept is a non-local transition. Recall Fodor’s chief complaint about P2.

¹⁵ Strictly speaking Fodor holds that context sensitive processes are either local or informationally intractable because of what he calls “the globality problem”: The problem that potentially all of one’s beliefs are relevant for the inferences implemented by such processes.

¹⁶ Fodor points to simplicity and relevance as examples of metrics which—given some candidate explanations—could be used to determine which explanation is the best explanation. But simplicity and relevance are non-local context dependent properties. The simplicity or relevance of a belief cannot be evaluated simply in terms of its local properties such as constituent structure. (*LOT2*, Ch. 4)

He claims: Because there are so many different kinds of relations in which a concept and its stereotype may stand, we cannot “just look at a stereotype and figure out what concept it is the stereotype of...” (*LOT2*, pp. 154-155) Thus the relevant information for determining the transition from a stereotype to concept acquisition requires non-local information. That is: Information that is not already in the constituent structure of the stereotype. Moreover, Fodor’s claim that there is no general inference pattern, by which given some stereotype identifies a concept, suggests that the inference is context sensitive. So by Fodor’s lights it seems that P2 cannot be a computational process. But given how badly CTM fares in accounting for processes that are undoubtedly in the domain of cognitive psychology and admit of psychological explanation (e.g., abductive reasoning), not being a computational process by the lights of Fodor is a far cry from being a merely brute process that must be explained at the neuro-biological level.

Recall that earlier we saw that Fodor treats inferential-rational processes and brute-causal processes as exhaustive categories. He makes the same move with respect to computational processes as well. That move appears as early as the last chapter of *LOT1*. A close reading of the final pages shows that the “new radical conclusion” of *LOT2* is already emerging. Thus Fodor’s new conclusion appears as a prescient worry in his original formulation of CTM. Here are just three noteworthy passages:

“The intended claim is that the sequence of events that causally determines the mental state of an organism will be describable as a sequence of steps in a derivation if it is describable in the vocabulary of psychology at all...” (*LOT1*, p. 198)

“Cognitive explanation requires not only causally interrelated mental states, but also mental states whose causal relations respect the semantic relations that hold between formulae in the internal representational system. The present point is that there may well be mental states whose etiology is precluded from cognitive explanation because they are related to their causes in ways that satisfy the first condition but do not satisfy the second.” (*LOT1*, p. 202)

“I think it’s likely that there are quite a lot of kinds of examples of causal but noncomputational relations between mental states. ...If this hunch is right, then these are bona fide examples of causal relations between mental states which, nevertheless, fall outside the domain of (cognitive) psychological explanation. What the cognitive psychologist can do, of course, is to specify the states that are so related and say that they are so related. But, from the psychological point of view, the existence of such relations is simply a matter of brute fact; explaining them is left to lower-level (probably biological) investigation.” (*LOT1*, p. 203)

These passages exemplify Fodor’s quick move from being non-computational to being brute.

1.5 Conclusions and reflections

This chapter has been dedicated to tracing and analyzing Fodor’s path to what I have been calling his “new radical conclusion.” The claim that concept acquisition is not a cognitive process, is not in the domain of cognitive psychology, does not admit of psychological explanation, and cannot be explained at the intentional level. I reviewed six variations on an argument for the new radical conclusion. I suggested that Fodor uses the underdetermination of concepts by stereotypes to motivate at least four versions of that argument (versions (c)-(f) above).

For the remainder of the discussion I will focus on the last two conjuncts of the new radical conclusion: That concept acquisition does not admit of psychological explanation and cannot be explained at the intentional level. I am limiting the discussion in this way because I take those to be the most important and most radical of the four conjuncts that make up the new radical conclusion. Depending on one’s preferred conditions for being a cognitive process, what is cognitive may turn out to be a small number of mental processes. And if being cognitive is a necessary condition of being in the domain of cognitive psychology, then what is in the domain of cognitive psychology will also turn out to be a small number of mental processes. What matters is the number of mental processes that we think do admit of psychological explanation and are explained at the intentional level, because that *is* a large number of mental processes. At least, that is the pervasive view both in empirical psychology and in philosophy.

Recall that the six specifications of NRC correspond to six properties. Each argument treats its corresponding property as a condition on concept acquisition being a process that admits of psychological explanation and is explained at the intentional level. So in effect each property carries an implication for demarcating the domain of processes that admit of psychological explanation and are explained at the intentional level. Since I read Fodor as holding the new radical conclusion for each of its conjuncts given any of the properties (modulo (a) and (b)), narrowing our focus should not change the relevance of any of the preceding discussion.

I reject each of the six arguments for the new radical conclusion because the six properties (as characterized by Fodor) are neither a necessary nor sufficient condition for concept acquisition being a process that is subject to psychological explanation and explained at the intentional level. I will not discuss (a) and (b)—the properties of being a process of hypothesis formation and confirmation, and being a learning process since, as I mentioned earlier, most of the dialectic surrounding Fodor's old and revised arguments have centered on these properties. In the case of (c) and (d)—the properties of being a computational process and a process that preserves truth and/or semantic content—my rejection of the argument turns on Fodor's own characterization of computation. (Fodor uses (d) and (f) at least in part to demarcate the domain of processes that are computational).

As I mentioned in section 1.4.5, Fodor repeatedly laments that CTM cannot explain many processes that are a matter of psychological interest and that we think should be subject to psychological explanation. (See especially, *LOT1*, pp. 197-205; 2000; *LOT2*, pp. 101-126) In particular Fodor has repeatedly noted that CTM's failure to explain perception and abduction show that there is something seriously wrong with CTM (2000, *LOT2*). By Fodor's lights, that

CTM cannot explain perception and abduction is a mark against CTM. By contrast, in the case of concept acquisition it seems that Fodor uses the fact that concept acquisition cannot be explained by CTM to argue for the new radical conclusion. Such comments go at least part of the way towards an admission that his notion of computational processes cannot be used to demarcate the boundaries of processes that are subject to psychological explanation.

In the case of (e)—the property of being an evidential process—my criticism turns in part on what I take to be Fodor’s motivation for considering the property in the first place. Fodor’s invocation of epistemic underdetermination derives from the more general idea that there seems to be an importantly epistemic aspect to the relationship between a newly acquired concept and the experience which elicits the acquisition of that new concept. I wholeheartedly endorse that general idea, and discuss it in detail in Chapter 2. What I reject is Fodor’s insistence that the relevant epistemic relationship is that of being evidential. Moreover, it is implausible as a condition for demarcating the domain of processes that admit of psychological explanation.

An upshot of disentangling the six versions of the argument is that—independently of the new radical conclusion—it reveals six different questions we can ask about the process by which a new concept is acquired. Namely,

- (1) Is concept acquisition a process of hypothesis formation and confirmation?
 - (2) Is concept acquisition a learning process?
 - (3) Is concept acquisition a computational process?
 - (4) Is concept acquisition a process that preserves truth and/or semantic content?
 - (5) Is concept acquisition a process in which some experience is treated as evidential for the acquired concept?
- And,
- (6) Is concept acquisition an inferential/rational process?

As I noted previously, most of the dialectic surrounding Fodor’s writing on concept acquisition has focused on the first two of these questions. Addressing the third question would require a thorough explication of the notion of “computational processes.” That is an interesting project

but one that is outside the scope of this dissertation. The use of the term covers a myriad of processes throughout the various disciplines housed under the umbrella of cognitive science. Some of these processes are described at the neural level and some are described in intentional terms at the level of psychology.

The answer to the fourth question is quite clearly “no,” for reasons that I discussed in section 1.4.4. Specifically, acquiring a new concept is an increase in expressive power, meaning that the representational repertoire of a mind that acquires some new concept cannot be characterized using the representational resources before the acquisition of the new concept. And, as I explained in the same section, it is difficult to come up with a construal of how a concept acquisition process could be said to “preserve truth.”

I take the answers to the fifth and sixth questions to be “no” as well, but the questions do suggest the more general and far more interesting question I gestured at above: Is the process of concept acquisition epistemically evaluable? That is, can the process be an instance of doing well by epistemic standards? This question is taken up in following chapter.

Chapter 2

How to Acquire a Concept Well (Epistemically Speaking)

Introduction

Consider the following question: Can we be warranted in acquiring a concept? This question might seem strange. Epistemic notions such as warrant are usually applied to beliefs and belief forming processes. Beliefs have the representational function of being true. Belief forming processes have the representational function of producing true beliefs. But unlike beliefs, concepts are not evaluable for truth or falsity. For example, the concept CAT¹⁷ cannot be true or false. Independently of some application, the concept CAT is neither accurate nor inaccurate. I grant all that. I want to ask whether we can epistemically evaluate the process by which a concept is acquired. That is, I am in part asking: Can an individual (or her psychology) do well—epistemically speaking—in acquiring a concept?

The idea that concept acquisition is epistemically evaluable is most intuitive in cases where some concept, C, is acquired on the basis of perceptual experience with a particular, where that particular instantiates the property indicated by C. For example, when a child acquires the concept DOG on the basis of perceptual experiences of dogs, we have the intuition that things are as they should be—that the relevant cognitive mechanisms are functioning well, relative to the child's representational capacities. Had she somehow acquired the concept DOG on the basis of perceptual experiences of cats we would judge that something had gone wrong with respect to

¹⁷ I use the convention of small caps to indicate concepts. As I use the term 'concepts', concepts are representational contents. Representational contents are abstract. They are psychological kinds that play a role in psychological explanation since mental states are typed in part by their representational contents. Thus my usage of the term 'concept' is closer to the Fregean notion of *sense* than the Representational Theory of Mind's (RTM's) notion of a concept. Everything that happens in this paper should be compatible with the view that concepts are mental representations or symbols. I will at times use the common locution 'the representational content of the concept C'. I use that locution for the sake of continuity with the RTM usage of the term which dominates the literature.

the child's representational capacities. These intuitions suggest that there are some normative standards by which we judge the process of acquiring a concept on the basis of perceptual experience. I will argue that at least some of these standards are distinctively epistemic.

The goal of this chapter is two-fold: First, I want to motivate the claim that concept acquisition is epistemically evaluable, and thereby show that we need an epistemology of concept acquisition. Second, I will introduce and defend a type of warrant for concept acquisition I call "concept formation entitlement."

My discussion will be organized around what I call "the evaluability question" and "the epistemic challenge". They are formulated as follows:

The Evaluability Question: Is concept formation subject to epistemic normative standards?

The Epistemic Challenge: Under what conditions is concept formation warranted?

As I use the expression 'epistemic warrant,' epistemic warrant is an epistemic good that is achieved by a psychology's meeting norms or standards that contribute to the fulfillment of the representational function of forming or preserving true beliefs.¹⁸ To say that "S is epistemically warranted in acquiring or forming some concept C by process *M*" entails (at least) the following: In acquiring C by *M*, S (or S's psychology) has done well with respect to norms that govern good routes to fulfilling the representational function of forming or preserving true beliefs. Ultimately I will argue that being warranted in acquiring C by *M* yields a prima facie warrant to employ C in forming beliefs.

¹⁸ My use of the term 'warrant' differs from the common usage often attributed to Plantinga. On Plantinga's usage 'warrant' is "that property or quantity that distinguishes knowledge from mere belief" but which is distinct from justification. (Plantinga, 1995) My usage follows Tyler Burge's taxonomy and teleological framework for epistemic kinds. (1993, 2013, pp. 3–4) However, Burge does not address the question of whether warrant applies to concept acquisition. Some people use the term 'justification' for the broad epistemic kind I am calling 'warrant'. For example see Pryor's usage of the expression 'having justification' in Pryor (2001, fn. 32, 2005, p. 182). On my usage, following Burge, epistemic entitlement and justification are two *species* of warrant. I explicate the distinction between the two in section 4.1 below.

The evaluability question asks whether a concept acquisition process can be evaluated against some standards for achieving an epistemic good such as warrant. The epistemic challenge presupposes an affirmative answer to the evaluability question. The epistemic challenge demands an account of the conditions under which some concept formation process can go well by the epistemic standards for achieving warrant. Distinguishing between epistemic evaluability and conditions on epistemic warrant makes it easier to see the full logical space of possible positions. For example, even if I convince you that concept acquisition is an epistemically evaluable process, you might disagree with me over the conditions under which a concept formation process can yield warrant.

In the present chapter I restrict the scope of the discussion to natural kind concepts that indicate properties instantiated in observable particulars such as CAT, KOALA, LEMON, etc.¹⁹ Nevertheless, explicating and answering the evaluability question and the epistemic challenge with respect to this restricted class of natural kind concepts will take us a good part of the way towards motivating an epistemology of concept acquisition. The chapter will proceed as follows: In sections 1 and 2 I lay out considerations in favor of an affirmative answer to the evaluability question. In section 3 I introduce Eric Margolis's "essentialist model" for natural kind concept acquisition. In section 4 I reintroduce the epistemic challenge and direct it at the essentialist model. I consider an evidentialist and a reliabilist condition for warranted concept acquisition. I reject the former and argue that latter must be supplemented. In section 5 I introduce my account

¹⁹ Specifically, the account does not apply to natural kind concepts that indicate properties instantiated in particulars that cannot be observed such as ELECTRON or QUARK. Ultimately, my account of what I call "concept formation entitlement" applies to a much larger class of concepts than the restricted class of natural kind concepts discussed in this chapter. I describe the larger class in Chapter 3 of this dissertation. The account is most intuitive when introduced in terms of natural kind concept acquisition.

of concept formation entitlement. Finally, in sections 6 and 7 I consider an objection to the epistemic evaluability of concept formation and offer some concluding remarks.

2.1 Motivating an affirmative answer to the evaluability question

While the epistemic evaluability of concept acquisition is rarely explicitly addressed in contemporary philosophy, the project connects with two well-established debates at the intersection of philosophy of mind and epistemology. The first is the debate over the possibility of concept learning. The contemporary dialectic of this debate is best characterized as a dispute over the nature of the process(es) by which concepts are acquired.²⁰ The second debate concerns the epistemic status of beliefs based on perceptual states. A central issue in this debate is whether and how perceptual states could provide epistemic support.²¹ Because the concept formation processes I consider are all based in part on perceptual experience, many of the worries that arise in the debate over the epistemic status of perceptual beliefs arise when considering the epistemic contribution made by concepts formed on the basis of perceptual states.

To motivate the project I will begin by explaining the connection between natural kind concepts and the objective of achieving empirical knowledge. Then I will show that there is an important gap in our epistemological theories of empirical knowledge, or at least in the epistemological theories of warranted empirical belief.

2.1.1 Forming natural kind concepts in pursuit of knowledge

I will begin by situating natural kind concept acquisition in the larger epistemic project that is the pursuit of knowledge of the empirical world.

²⁰ See for example, Fodor 1975, 1981, 2008, Carey 2009, Samules 2002, Prinz 2002, and Rey 2014. The contemporary debate evolved from early modern disputes between rationalists and empiricists over innate ideas and the richness of innate concept acquisition machinery.

²¹ For just a few examples see Sellars 1956, Burge 2003, Pryor 2005, Schellenberg 2014, and Silins 2011. (Strictly speaking, Sellars (1956) did not recognize perceptual states, and wrote in terms of sensations epistemically supporting beliefs).

Concepts are constitutive representational elements of propositional attitudes. Though natural kind concepts can occur in various propositional attitude types, I claim that they are constitutively associated with the formation of *empirical beliefs*. An empirical belief is a belief whose warrant is at least in part dependent for its force on sense experience. Natural kind concepts are constitutively associated with empirical belief formation because their representational function is to indicate and attribute observer-independent properties that are the subject matter of empirical beliefs. Nearly all empirical beliefs are beliefs that have the objective observable world as their subject matter. (Hereafter when I use the term ‘world’, that is the restricted notion I will be employing).²² All belief formation systems have the *representational* function of forming true beliefs.²³ Natural kind concept formation serves empirical belief formation by producing concepts that make a contribution towards forming true empirical beliefs.

I propose to think of empirical belief formation and natural kind concept formation as part of the greater epistemic endeavor I call “empirical discovery”. Empirical discovery is (at least in part) the process of forming both natural kind concepts and empirical beliefs. Think of the world as being described by a set of truths. The epistemic aim of empirical discovery is to learn empirical truths—and ultimately—to achieve an understanding of the nature and structure of the empirical world. Discovering natural kinds is a part of this epistemic endeavor. One does

²² The belief that, IF FIDO IS A DOG, THEN FIDO IS A DOG, does not qualify as an empirical belief despite having some part of the objective observable world as part of its subject matter. I write “nearly all” because beliefs about one’s own pains do not meet the second condition. Beliefs about one’s own pains are empirical because their warrant is at least in part dependent for its force on sense experience but their subject matter is neither observable nor objective. Thus I take the former condition—that the warrant for empirical beliefs is at least in part dependent on sense experience—to be the fundamental condition for being an empirical belief.

²³ It is important not to conflate the notion of representational function with teleological accounts of mental states that appeal to biological functions. For an extended discussion of various functions see Burge 2003, pp. 509-11, 2010a, pp. 291-315.

not achieve understanding of the structure of the world by simply collecting truths about how a variety of particulars have behaved under various conditions at some time. The challenge of discovery is that to learn empirical truths is not merely to compose and confirm a list of hypotheses and attendant theories. A system that engages in discovery must create the representational tools from which such hypotheses and theories are composed. Natural kind concepts are some of these representational tools. So forming natural kind concepts is part of the greater epistemic pursuit of empirical knowledge.²⁴

Once we think about the total epistemic endeavor that is the pursuit of empirical knowledge, we can see that there is an important gap in our epistemology. To bring this gap into light we need to look at the pursuit of empirical knowledge “from the ground up” so to speak.

For illustration, consider the pursuit of empirical knowledge about tomatoes. Suppose that no one has ever seen or heard of tomatoes. You are exploring some previously unknown island and lo and behold you come across a tomato vine. Hanging on the vine is a beautiful red ripe tomato. You form a perceptual state (*P-SI*) with the content ***that red round body there***.²⁵ Currently, epistemology has some accounts of why you are warranted in transitioning from this perceptual state to the perceptual belief (B1) THAT IS A RED ROUND BODY THERE.²⁶ But importantly, our pursuit of empirical knowledge is more than just the pursuit of singular beliefs

²⁴ My thinking of natural kind concepts as representational tools grew out of discussions with Mark Greenberg and his 2012 seminar on concepts.

²⁵ I use asterisks and italics as a convention to indicate the representational content of perceptual states. I am assuming a non-conceptualist but intentional account of the representational content of perceptual states. On this account to represent something *as* red and round in perception does not require having the concepts RED or ROUND.

²⁶ The perceptual representational content and its conceptualized counterpart differ only in the form of the state. Specifically, perceptual beliefs are propositional. The characterization of the representational content of (*PS-I*) is a gross simplification. For example, I have left out the demonstrative elements guided by each perceptual attributive. For a detailed discussion of the representational content of perceptual states and the distinction between perceptual and conceptual representational content see Burge 2003, 2010a, 2010b.

about instances of perceptually discriminable properties. Recall that empirical discovery aims at achieving an understanding of the nature and structure of the empirical world. Achieving this requires general empirical beliefs and inductively rich categories. Red-round-bodies is not an inductively rich category because there aren't many explanatory generalizations about all the red round bodies out there. To get greater knowledge and understanding of the nature and structure of the empirical world we need to be able to represent the world in ways that go beyond the ways we represent the world in perception. In order to do that we need to form concepts that indicate properties that are not represented in perception.²⁷

Let us go back to you, the expert botanist on the island, and that red round body hanging on the vine. You cannot represent that red round body *as* a tomato before you form the concept TOMATO. Once you have the concept TOMATO, our epistemological theories have a lot to say about how you should go about forming warranted beliefs about tomatoes, how you should adjust your tomato credences, and how to best pursue your future tomato research. So the current state of our epistemological theories of empirical knowledge is the following: there are theories about standards for transitioning from perceptual states to perceptual beliefs, and there are theories about standards for forming beliefs and adjusting your credences. But our epistemological theories are silent on standards for transitioning from having perceptual beliefs about red round bodies to having beliefs about tomatoes, in cases where that transition requires the formation of the concept TOMATO.

The point of the story about the expert botanist discovering tomatoes is to highlight the following: our warrant for forming beliefs about empirical subject matter of which we know almost nothing, is inextricably linked to our ability to form new ways of representing the world.

²⁷ The scope of this paper prevents me from arguing for the claim that there is no perceptual attributive that represents entities *as* tomatoes. For the purposes of this paper I will just assume that this is true.

The expert and the infant are in much the same boat in this respect. Let us call empirical subject matter about which we know almost nothing, “the unknown”. Furthermore, let us call the first beliefs that are formed about some subject matter “early beliefs.”

Empirical beliefs carry empirical commitments about the nature of the world. The expert and the infant are in much the same boat with respect to forming empirical beliefs about the unknown. That is the case because we can have concepts that indicate properties about which we know very little. That this is so is most salient in the case of early beliefs. What we know might not be sufficient for successful application of a newly acquired concept in many conditions, and our explications of a newly acquired concept can fall far short of an explication sufficient to individuate that concept. Moreover, though one cannot tell definitively by observation that something is in fact an instance of some natural kind, one can still be warranted in their belief that it is.²⁸ Still, in the face of this fallibility and limited knowledge form early empirical beliefs about the unknown that employ newly formed natural kind concepts and thereby make empirical commitments about what it is to be the kind of thing that has the property indicated by the newly acquired concept. The gap in our epistemological theories concerns the source of the warrant for forming early beliefs about the unknown that employ newly acquired concepts.²⁹

²⁸ As a series of seminal papers by Hilary Putnam and Tyler Burge suggest, there is no shortage of examples where a subject, S, is grossly ignorant of true facts about some property such as the property of being a tomato, and yet it seems appropriate to ascribe to S mental states with the representational content that includes the concept TOMATO. It is consistent with these thought experiments that there may be some inferential relation(s) that is/are constitutive of some concept. But these relations may not be sufficient to individuate a concept. And even if there are inferential relations that are part of the individuation of some concept, it is a further substantive claim that those individuation conditions will determine a concept's possession conditions. For a general criticism of treating individuation conditions as possession conditions see Greenberg (2013). For the original thought experiments see Putnam 1962/1975a, 1970/1975c, 1973/1975b, 1975d and Burge 1979/2007a, 1982/2007c, 1986/2007b. I plan to discuss these issues in more detail in future work.

²⁹ At this point the following objection might be raised: Early beliefs about the unknown that employ newly formed concepts can be warranted through a general disposition to defer to experts as the final arbiters of our empirical beliefs. According to the deference objection, the infant (among others) is entitled to early beliefs such as THAT IS A CAT, because on the matter of cats she is disposed to defer to cat experts, or to nearby adults who are themselves disposed to defer to cat experts. On such a view the process by which concepts are formed plays no role in the

I am arguing that our warrant for forming beliefs about the unknown is underwritten at least in part by our ability to form new ways of representing; and that this only makes sense under certain conditions. My claim is that those conditions are about the processes by which we form new ways of representing the world—that is—the processes by which we form our concepts. And these are the processes that we need in order to explain the epistemic continuity from perceiving the world to achieving empirical knowledge. Concept formation processes are the bridge between perceptual states/beliefs and early belief formation.

2.2 The epistemic contribution of concept formation processes

The discussion so far has been about locating the place of natural kind concept acquisition processes in the larger epistemic endeavor I called empirical discovery. Before proceeding any further we must clearly distinguish between *having* some concept C and the *process* by which C is formed. The idea that there is a difference between a process of forming x, and x itself, is pretty straightforward. The process of forming or acquiring some concept C is the process that eventuates in the having of C. Having some concept C is having a representational competence such that you can entertain a propositional attitude whose representational content includes C. So, for example, if S has the concept DOG, then S can entertain the representational content DOG. When I speak of forming a concept, I am talking about the first time it is created, which—in the normal case—is the first and only time the concept is created within a given subject's psychology.³⁰ Thus forming a concept is different from employing a concept. A

explanation of our warrant to form beliefs that employ newly formed concepts. I cannot address this objection in full here, but I am skeptical that appeals to deference can account for all cases. There are plausible cases in which an individual is not in a position to defer to experts, because there are no experts about the subject matter in question. But more importantly, the deference objection misplaces the ultimate source of our warrant. I plan to develop both the deference objection and my reply in future work.

³⁰ Non-standard cases include disease or injury that could result in damaging or destroying some conceptual competence but without injuring the concept acquisition mechanism, making it possible to acquire the concept for a second time. A different kind of case in which a concept could in principle be acquired more than once within a

concept is a representational competence that can be employed many times in the formation of various beliefs and other propositional attitudes.

Keeping the distinction between having a concept and forming a concept in mind, consider the following four claims:

For some natural kind concept C and some process M by which C is formed

- (1) Having C can make a positive epistemic contribution to warranted empirical belief and knowledge.³¹
- (2) Forming natural kind concepts is part of the pursuit of warranted empirical knowledge.
- (3) M can make (or fail to make) a positive epistemic contribution to warranted empirical belief and knowledge.
- (4) M is epistemically evaluable.

Notice that (1) does not entail (2), (3) or (4). Nor does (2) entail (1), (3), or (4). (1) is a claim about *having* the representational competence typed by the concept C . (2) is a claim about which processes are part of the pursuit of empirical knowledge. Claims (3) and (4) are about the epistemic import of the process by which C is formed. Someone might agree that (1) and (2) are true, but then rightly point out that the truth of (1) and (2) does not entail the truth of (3) or (4).

My goal is to argue for (3). In the rest of this section I will propose what I take the positive epistemic contribution that concept formation processes can make to warranted empirical belief and knowledge. Then I will explain why if (3) is true then (4) is true.

My proposal is the following: the process by which some concept C is formed determines whether an individual (or her psychology) has prima facie warrant to employ C in the formation of beliefs about the subject matter indicated by C . Reformulating the proposal specifically for the

psychology is if the concept is acquired first in some subpersonal informationally encapsulated module that is inaccessible to central systems or other modules, and then acquired again in some other module.

³¹ At least in worlds where C indicates a property that is instantiated in that world at some time.

case of natural kind concepts and empirical beliefs we get: the process by which some natural kind concept *C* is formed determines whether an individual (or her psychology) has prima facie warrant to form empirical beliefs about the subject matter indicated by *C*. That prima facie warrant is the positive epistemic contribution that the process of forming a concept can make to empirical knowledge. Merely having a concept does not come with any such prima facie warrant. If this is right then only some cases of having a concept come with prima facie warrant to employ that concept in the formation of beliefs.

How does the above bear on the evaluability question? Recall claims (3) and (4) presented above:

- (3) *M* can make (or fail to make) a positive epistemic contribution to warranted empirical belief and knowledge.
- (4) *M* is epistemically evaluable.

Earlier I said that if (3) is true then (4) is true. That is, if some process *M* by which a concept is formed can make (or fail to make) a positive epistemic contribution to empirical knowledge, then *M* is epistemically evaluable. Thus far I have discussed one way in which the process can make a positive epistemic contribution. To make a positive epistemic contribution, the process must produce concepts in a way that is non-accidental relative to our representational capacities. The epistemic evaluation of a concept formation process just is the evaluation of whether or not *M* makes a positive epistemic contribution to warranted empirical belief and knowledge. If this is right, then the process by which a concept is acquired has epistemic import and thus can be evaluated by epistemic standards.³²

³² I do not assume that *all* cases of natural kind concept acquisition are dependent on direct experience of a particular that instantiates the property indicated by the acquired concept. It seems clear that one need not have an experience of an instance of some natural kind *x* in order to acquire the natural kind concept *x*. For example, the concept OCTOPUS can be acquired by being told a story about an octopus or by being shown pictures of an octopus. My account of concept formation entitlement for natural kind concepts also applies to cases in which natural kind

2.3 A Model for acquiring natural kind concepts

Sections 1 and 2 defended the claim that a concept acquisition process is epistemically evaluable. In section 4 I will turn to the epistemic challenge and ask under what conditions is concept formation warranted? In order to answer that question we will need a candidate model of concept acquisition. To this end I will employ a rough caricature of Eric Margolis's model for the acquisition of natural kind concepts (Margolis, 1998).³³ I call this rough caricature "the essentialist model".

2.3.1 The Essentialist Model

Margolis suggests thinking of a concept as a file, where the file label specifies which concept it is, and the entries in the file are knowledge structures associated with the concept. But these entries do not determine the content of the concept. Margolis adopts Jerry Fodor's information-based atomistic theory of concepts since his essentialist model is a response to Fodor's claim that all concepts are innate.³⁴ On Fodor's version of information-based atomism, a concept has its content in virtue of getting nomologically locked to a property in the world, thereby coming to carry information about that property. Once a concept is locked to a property in the right way, the concept indicates that property. That is, a concept C has the content CAT

concepts are formed in these other ways. I believe that the explanations of these other cases are ultimately parasitic on the explanation of the direct experience case.

³³ Margolis's model has garnered support among both philosophers and psychologists. For example, Susan Carey (2009) endorses the general contours of Margolis's model. And Strevens (2012) builds on Margolis's model to motivate a non-inferential psychological process he calls "introjection." The essentialist model draws on the psychological essentialist literature in developmental and cognitive psychology. Importantly similar models have been put forward by Macnamera 1986, Keil 1989, Oved 2009, and arguably in Strevens 2000, as well as others. Margolis's model is developed and defended further in Laurence & Margolis 2002 and Margolis & Laurence 2011.

³⁴ Strictly speaking Fodor's (1975) claim is that all *monomorphemic* concepts are innate. Monomorphemic concepts are concepts that correspond to morphemes, typically characterized as the smallest meaningful units of a natural language. Concepts are either monomorphemic or complex. Complex concepts are concepts that decompose into monomorphemic concepts. For example, the concept BLACK CAT is a complex concept since it can be decomposed into two smaller meaningful units—the monomorphemic concepts BLACK and CAT.

because C is lawfully and reliably correlated (in the right way) with the property of being a cat.³⁵

One of the supposed upshots of atomism is that the possession of some concept C does not depend on the possession of any particular knowledge structure. So, for example, my possession of the concept CAT does not depend on my having any particular belief about cats.

It is difficult for information-based atomism to explain the non-arbitrary relation between the experience of some property and the acquisition of the concept that indicates that property. Because the possession of some concept C is cashed out simply in terms of a nomological locking to the property indicated by C, the process by which the locking is achieved is at least in principle independent of the identity of the concept acquired. Nothing in Fodor's theory requires that the locking be achieved by a process that originates in some kind of causal interaction with the property indicated by C.³⁶ To address this difficulty Margolis introduces a type of sustaining mechanism for natural kind concepts. On his proposal, two types of knowledge structures explain the acquisition of natural kind concepts on the basis of perceptual experience. They are: a kind syndrome and an essentialist assumption. I will explain each of these in turn.

A kind syndrome is "a collection of salient properties that are readily open to inspection and are reliable indicators that something is of that kind" (Margolis, 1998, p.356). When talking

³⁵ The "right way" on Fodor's model involves an asymmetric dependence relation among the laws that hold between concepts and properties. I leave out the details of asymmetric dependence since the details are not immediately relevant to the present discussion.

³⁶ Fodor (1998, 2008) acknowledges that his view has a problem accounting for why it is so often that a concept is generally acquired from experience with good instances of the property that the concept indicates. He calls this the "doorknob/DOORKNOB problem." Fodor opts for a metaphysical solution to the d/D problem. According to Fodor the concept DOORKNOB tends to be acquired from good instances of doorknobhood because part of what it is to be a doorknob is to cause the acquisition of the concept. He writes, "[I]t used to seem to me that atomism about concepts means that DOORKNOB is innate. But now I think that you can trade a certain amount of innateness for a certain amount of mind-dependence. Being a doorknob is just: striking our kinds of minds the way that doorknobs do. So, what you need to acquire the concept DOORKNOB "from experience" is just: the kind of mind that experience causes to be struck that way by doorknobs." (1998, p. 162) For present purposes I take this solution to the d/D problem to be a non-starter. I discuss it in more detail elsewhere in my dissertation.

about a *representation* of a kind syndrome I will use the term ‘stereotype,’ though Margolis does not employ this term. A stereotype is a perceptually derived statistical representation of the co-instantiation of the salient properties that make that are reliable indicators that something is of that kind. A stereotype is a statistical representation because in addition to representing perceptually discriminable features, a stereotype includes some indication that the features that make up a kind syndrome are frequently co-instantiated in the world.^{37, 38} I will follow Margolis and use his example of the stereotypical salient observable features of cats. These features include: characteristic shape, characteristic colorings, and texture (such as being grayish-brown and being furry), characteristic motions including gait, and characteristic sounds made by cats.³⁹ A cat stereotype is just a representation of these statistically co-instantiated features, along with some indication that the features tend to be co-instantiated. For most of what follows I will talk about acquiring concepts on the basis of the representational content of stereotypes rather than perceptual states.

A representation of a kind syndrome is not constitutive of a concept’s content. Rather, a syndrome is a collection of features that are *symptomatic* of the property that a concept indicates. Representations of natural kind syndromes have the potential to explain the non-arbitrary relation

³⁷ I believe that the notion of a stereotype in the text is the same notion that Fodor has in mind in what he calls his “story about concept acquisition” (2008, p. 162) discussed in Chapter 1. The idea that learning a stereotype is a way of learning the meaning of natural kind term is developed in Putnam’s early work on semantics (1970/1975c).

³⁸ Such statistical representations are sometimes described as representations of “feature clusters” within a feature space, where clustering occurs in virtue of some similarity and dissimilarity functions. There are various models of statistical representations in the perceptual, developmental, and learning literatures. Many of these are geometric models of similarity using multi-dimensional scaling (Goldstone and Son, 2005). Alternative set-theoretical approaches to modeling similarity have also been proposed (Tversky, 1977). I discuss feature spaces and similarity models in more detail in Chapter 3.

³⁹ Margolis draws on the experimental literature on concept acquisition to motivate these features. In particular he considers the psychological literature on perceptual motion cues, and shape bias. The scope of this paper prevents me from giving a review of the experimental literature that Margolis cites to support the claim that shape and motion as good candidates for “salient” properties.

that an experience of a natural kind has to the acquisition of a natural kind concept. Natural kinds tend to have reliably correlated superficial properties, though the particular set of superficial properties is not constitutive of being a member of that natural kind. Importantly, experience of the co-instantiation of these superficial non-constitutive properties is generally a good guide to tracking kinds. When talking about a representation of a kind syndrome I will continue to use the term ‘stereotype’, though Margolis does not employ this term.

2.3.2 The projection problem revisited: stereotypes underdetermine concepts

Notice that the acquisition of a concept on the basis of a stereotype is an instance of a transition between psychological representations, where the representational content of the first representation underdetermines the representational content of the second. In Chapter 1 I called the underdetermination of concepts by stereotypes “the projection problem,” and provided the following example: the cat stereotype underdetermines the representational content of CAT, because one could project to a multitude of different concepts based on the representation of the statistically co-instantiated features included in the cat stereotype. Let us call the stereotypical cat features I mentioned earlier (the characteristic shape, coloring, texture, motion pattern, and sound) *features 1-5*. Features 1-5 are also stereotypical of the entities in the extension of the concept ANY THING THAT EXHIBITS 1-5. Call this latter concept the SHMAT concept. The concept CAT and the concept SHMAT will have different extensions. Abnormally looking and abnormally moving cats are included in the extension of CAT. However, these unfortunate souls may be too deviant to make it into the extension of SHMAT. Moreover, cat-fakes—(non-cats that exhibit features 1-5)—will be in the extension of SHMAT. Although they have different extensions, CAT and SHMAT can share the same stereotype. Putnam’s famous scenario in which all of the entities on Earth that we think of as cats actually turn out to be robots controlled by radio from Mars

(hereafter simply “robocats”) provides yet another example of a kind that would produce the same stereotype as the kind cat (Putnam, 1962).⁴⁰

One way to think of the projection problem is as follows: Given what is included in a stereotype, why should the cat stereotype lead a child to acquire the concept CAT as opposed to the concept SHMAT? The preceding examples show that concepts with different extensions can share a stereotype. But the projection problem also arises in cases of coextensive but intentionally distinct concepts. To use a variation on one of Fodor’s favorite examples, suppose cats are your grandmother’s favorite animal. The cat stereotype will also be the stereotype produced by instances of the property of being Granny’s favorite animal. To overcome underdetermination, we need an explanation of why the cat stereotype leads to the acquisition of CAT rather than the acquisition of GRANNY’S FAVORITE ANIMAL.

2.3.3 A solution to the projection problem

The second component of Margolis’s model provides one possible solution to the projection problem for natural kind concepts. Margolis describes this component as a “knowledge structure” that he calls “the essentialist assumption”. Margolis characterizes the essentialist assumption as a belief that:

The Essentialist Assumption: Membership within the kind is determined by possession of an essential property (or set of properties) and this property is a reliable cause of the syndrome. (1998, pp. 357–8)

⁴⁰ Fodor uses what I am calling “the projection problem” to challenge the coherence of the idea of concept learning. In a recent paper on Susan Carey’s 2014 response to Fodor, Georges Rey suggests thinking of Fodor’s challenge as a version of Goodman’s riddle of induction (Rey, 2014). Though I see a connection between Rey’s discussion and my own, he is primarily concerned with matters outside the scope of this paper. It’s worth noting that—Fodor’s complaints notwithstanding—underdetermination in this case is not obviously problematic. Whether there are cases of acquiring multiple concepts on the basis of a stereotype is largely an empirical question. The claim that a single stereotype can lead to the acquisition of multiple concepts is noncommittal about whether the concepts are acquired at the same time. For example, perhaps the cat stereotype is sufficient for acquiring CAT and SHMAT, though not sufficient for acquiring ROBOCAT until some further conditions are met.

The essentialist assumption is drawn from the empirical literature on *psychological essentialism*. Psychological essentialism is a theory about the psychological tendency to essentialize categories. The central claim of psychological essentialism is that people (including children as young as 2-3 year olds) distinguish between the superficial properties and the essential unobservable deep structural properties that type-identify objects and are responsible for observable similarities that category members share. (Medin 1989, Medin and Ortony 1989, and Gelman 2003). The most robust empirical evidence of the disposition to essentialize has been found in the case of *basic-level* natural kind categories such as *cat, bird, dog, and tree*.⁴¹

The essentialist assumption as formulated by Margolis accounts for why the cat stereotype yields the concept CAT as opposed to the coextensive concept GRANNY'S FAVORITE ANIMAL, or the artifact concept ROBOCAT. The essentialist assumption reflects what we can call "natural kind grouping," which in conjunction with a stereotype yields a unique natural kind concept. Natural kind grouping privileges the transition from a cat stereotype to the natural kind concept CAT by treating the frequently co-instantiated salient features represented in the stereotype as merely symptomatic of a naturally occurring discoverable property, rather than treating these features as defining or constructing a category. A grouping that treated these features as criteria for group membership would yield the SHMAT concept described above.

⁴¹ An object is a member of a number of different hierarchically organized categories. For example, Ralphie (my parents' cat) is a member of the very general categories *object, living thing, and animal*, as well as more specific categories such as *maine coon, black male maine coon, and black male maine coon residing in Santa Cruz, California*. Murphy characterizes the basic-level category as "the most natural, preferred level at which to conceptually carve up the world." He suggests that the basic-level "can be seen as a compromise between the accuracy of classification at the maximally general level and the predictive power of a maximally specific level." (2002, p. 210) Note that such a characterization allows the "basic-level" to be located at different levels of abstraction relative to context. Examples of basic-level natural kinds are not universal across individuals because knowledge of a domain can affect the level of generality at which the basic-level is located. For an extensive review of the psychological literature on categorization and the basic-level see Murphy (2002, Ch.7). For a seminal paper motivating the idea of a "basic-level" see Rosch (1978). For some pioneering experiments on categorization see Rosch and Lloyd (1978). For a view that posits that some "global" concepts such as ANIMAL are developmentally prior see Mandler and McDonough (1993).

Thus the essentialist model explains the acquisition of a natural kind concept as follows: repeated experience with instances of a kind leads to the formation of a stereotype, which is a representation of a syndrome that is a reliable indicator of the kind. The perceptually derived stereotype, in conjunction with the essentialist assumption, sustains the mind-world relation between the new concept and the kind by generating the dispositions that Fodor's information-based theory requires. The concept is locked onto a natural kind property through a representation of the kind syndrome (i.e., the stereotype) along with the assumption that such syndromes reliably indicate natural kind properties. The stereotype along with the essentialist assumption are the knowledge structures that can be thought of as file entries on the concepts as files metaphor described above.

2.4 The Epistemic Challenge applied to the Essentialist Model

Now that we have the details of a concept acquisition process, we can ask what are the conditions under which a concept acquisition process would be warranted? In section 2 I claimed that our warrant for forming beliefs about that which we know almost nothing, is inextricably linked to our ability to form new ways of representing the world. I said that this link only makes sense under certain conditions. I claimed that those conditions are conditions on the processes by which we form our concepts. These are the conditions for achieving the prima facie warrant to employ newly formed concepts. In what follows when I speak of "a warranted concept acquisition process" or "a warranted concept formation process" I am referring to a process that meets the epistemic standards for achieving prima facie warrant to employ newly acquired concepts in the formation of beliefs. And specifically, the prima facie warrant for employing newly formed natural kind concepts in the formation of early empirical beliefs about the unknown.

In this section I presume an affirmative answer to the evaluability question and focus on the epistemic challenge. In section 4.1 I reintroduce the epistemic challenge and direct it specifically at the essentialist model. In section 4.2 I consider an evidential condition for warranted concept acquisition. In section 4.3 I consider a reliability condition. Ultimately I reject the evidential condition and argue that the reliability condition needs to be supplemented.

It is important to note that showing how a disposition for natural kind grouping in conjunction with a stereotype can yield a unique natural kind concept does not straightforwardly answer the epistemic challenge raised by the projection problem or the evaluative question. In order to address the epistemic challenge and the evaluative question we need an account of how it is that concept acquisition is subject to epistemic norms at all, and why natural kind grouping on the basis of the representational content of stereotypes reflects a “doing well” by epistemic standards. In the next section I begin addressing these issues by first reviewing the epistemic challenge and considering some candidate replies.

2.4.1 The Epistemic Challenge reintroduced

Recall that I formulated the epistemic challenge as follows:

The Epistemic Challenge: Under what conditions is concept formation warranted?

Thus far I have only discussed the notion of epistemic warrant in general. In the introduction I explained that epistemic warrant is an epistemic good that is achieved by a psychology’s meeting some standards for realizing the representational function of belief formation which is to form true beliefs. At this point it would be helpful to have some finer-grained epistemic notions. To that end I will use Tyler Burge’s distinction between justification and entitlement. (2013, pp. 3–4, 28) Entitlement and justification are species of warrant. Justification is warrant by reason. For an individual S, being warranted by justification consists at least in part in a reason being “operative or relied upon” in S’s psychology. Entitlement is warrant without reason. It is warrant

that derives purely by a psychology's operating well and reliably with respect to its representational functions.⁴² In contrast to warrant by justification, if S is warranted by entitlement, S does not need to represent (either personally or sub-personally) the propositional content that characterizes her warrant.⁴³ The epistemic challenge is formulated in terms of warrant so as to be neutral between these two species of warrant.

I presented the essentialist model so that we could consider the epistemic challenge in conjunction with a concrete example of a concept formation process. The essentialist model itself does not address the epistemic challenge. However the model does provides us with enough detail to test some candidate accounts of what *would* make acquiring a natural kind concept on the basis of a perceptual experience warranted. We can direct the epistemic challenge at the essentialist model by reformulating it as follows:

The Epistemic Challenge: Under what conditions is the formation of natural kind concepts on the basis of perceptually derived stereotypes warranted?

In the next two sections I will consider two competing conditions for warranted essentialist concept formation: an evidential condition and a reliability condition.

2.4.2 An Evidential Condition

In his many writings on the impossibility of concept learning, Jerry Fodor reflects on the relationship between an experience and a concept acquired on the basis of that experience. Fodor

⁴² Importantly, the relevant environment for judging the reliable operation of a psychological competence is the environment in which the competence is explained and established. I explain this aspect of Burge's use of reliability in section 4.4 below. For the original discussion in Burge, see especially, 2003, pp. 532–400.

⁴³ Reasons are propositional and must be represented in an individual's psychology. According to Burge, an individual S need not be able to access or bring a reason into consciousness to have warrant by justification. (2013) By contrast, S can have warrant by entitlement even if she could not acquire the concepts that would be necessary to think the propositional content that characterizes the nature of the warrant. (2003, pp. 504–5) Burge has developed the notion of entitlement with respect to self-knowledge, testimony, transitions in reasoning, memory-generated beliefs, various presuppositions, and the formation of perceptual beliefs (1993, 1996, 2003).

often appeals to the intuitive connection between evidence and rationality. In one such passage he writes,

“Perhaps COW is learned from experience with cows? If so, then experiences with cows must somehow witness that it’s cows that COW applies to. This internal connection between concept learning and epistemic notions like evidence is the source of the strong intuition that concept learning is some sort of rational process. It contrasts sharply with kinds of concept acquisition where, for example, a concept is acquired by surgical implantation or by swallowing a pill; or by hitting one’s head against a hard surface, etc.” (2008, p. 135)

Fodor emphasizes the “intuitive connection” between evidence and “rational processes”. I have not and will not be arguing for the claim that concept acquisition is a rational process. I take this to be a stronger claim than the claim that the process can be warranted. The terms ‘rational’ and ‘rationality’ are generally applied to processes that involve reasoning and inference, and thus connote a type of psychological process that is more intellectual than a warranted process needs to be.

Fodor’s meditation on the intuitive connection between evidence and rationality suggests the following condition for what *would* make acquiring a concept on the basis of experience warranted:

The Evidential Condition The acquisition of a concept C based on experience e is warranted if e is used as evidence for the property indicated by C .⁴⁴

According to the evidential condition, it is not sufficient that an individual have a disposition to acquire the natural kind concept CAT from the cat stereotype. The concept must be acquired on the basis of information that is used as evidence. So the cat stereotype would have to be used as evidence for the concept CAT.

⁴⁴ Here “used” should be read as “used by the individual or the psychology of the individual.” My formulation of the evidential condition is a charitable reconstruction of Fodor’s view. Fodor often writes as though one needs evidence for the conditions under which the concept CAT applies. See remarks by Fodor 2008, pp. 141–5 and 1998, p. 127, and his response to Stephen Laurence in an exchange that took place at the *Symposium on Solutions to Fodor’s Puzzle of Concept Acquisition* at the 2005 Cognitive Science Society meeting in Stresa, Italy. (Fodor et al. 2005)

One likely motivation for the evidential condition comes from an underlying worry about the epistemic consequences of what I referred to earlier as the “projection problem”. Recall that the projection problem is the semantic underdetermination of concepts by stereotypes. In discussing the projection problem I said that one way to resolve the projection problem is by appeal to Margolis’s essentialist assumption. The essentialist assumption secures the disposition to acquire the natural kind concept CAT from the cat stereotype. As I said above, Margolis characterizes the essentialist assumption as a belief that:

The Essentialist Assumption: Membership within the kind is determined by possession of an essential property (or set of properties) and this property is a reliable cause of the syndrome. (1998, pp. 357-8)

A proponent of the evidential condition would likely argue that if the essentialist assumption (or something like it) is *not* represented in the psychology of the individual forming a concept on the basis of a stereotype, then the warranting connection between a stereotype and a natural kind concept is severed. However if the essentialist assumption were represented in the psychology of the individual forming a concept, then it could provide a warranting bridge between stereotypes and the concepts that those stereotypes underdetermine. On such a view, the essentialist assumption would secure the essentialist disposition to form a natural kind concept on the basis of a stereotype by being deployed as a premise in a chain of reasoning. If such reasoning took place, then it might be plausible to construe the formation of a concept on the basis of a stereotype as a process that reflects a sensitivity to evidence. For example, such reasoning could employ the essentialist assumption as the belief that: A KIND IS DETERMINED BY POSSESSION OF AN ESSENTIAL PROPERTY (OR SET OF PROPERTIES) THAT IS A RELIABLE CAUSE OF A KIND SYNDROME; Or more simply, REPEATED EXPERIENCE OF A KIND SYNDROME IS EVIDENCE OF A NATURAL KIND.

A belief need not be conscious or even available for conscious reflection, but it must be represented, so the component concepts that make up the belief must be available at least subpersonally. Whether or not young children who manifest an essentialist disposition have such a belief is, of course, an empirical question. I am not optimistic that empirical data will support such an intellectualist construal of the essentialist assumption.⁴⁵ There is extensive empirical evidence that suggests that infants can acquire some kind concepts before 12 months of age (Carey 2009, p. 279). It seems improbable that these infants have concepts such as EXPERIENCE, EVIDENCE, and NATURAL KIND.^{46, 47} But more importantly, there is no need to construe the essentialist assumption as a belief in order to explain how infants form natural kind concepts. Similarly, we do not need to construe the essentialist assumption as a belief in order to explain how such a process could be warranted. The only motivation for doing so would be a theoretical commitment to the claim that warranted psychological transitions must be instances of rational inference. In the case of concept acquisition, warrant by entitlement is more plausible.⁴⁸

⁴⁵ Margolis's use of the term 'belief' in characterizing the essentialist assumption suggests this reading, but it is not clear that he is committed to such an intellectualist construal of the essentialist assumption. Simply converting the essentialist assumption into a premise in a chain of reasoning is an instance of what James Pryor calls the "premise principle." The premise principle states that, "The only things that can justify a belief that P are other states that assertively represent propositions, and those propositions have to be ones that could be used as premises in an argument for P. They have to stand in some kind of inferential relation to P: they have to imply it or inductively support it or something like that." (2005, p. 189) Pryor introduces the premise principle in a discussion of opponents to immediate justification. I think it is quite likely that Fodor's view assumes the premise principle.

⁴⁶ Limitations of space prevent me from making a case against the evidential condition based on this evidence. 12 months is a conservative, but well supported estimate. See for example Xu 2002, 2005, Xu et al. 2004, and Dewar and Xu 2007, and discussion of these papers in Carey 2009, pp. 266-76. Fodor's remarks (1998, pp. 157-8) suggest that he thinks that the concepts needed to state the essentialist assumption would have to be available (at the very least subpersonally).

⁴⁷ Fodor (1998) distinguishes having a natural kind concept from having a natural kind concept "as a natural kind". He claims that possession of the latter is a late intellectual achievement—both ontogenetically and historically—and requires also having concepts like microstructure, hidden essence and natural kind (1998, pp. 155-9). I think his argument is philosophically problematic and empirically unfounded. I discuss some of these issues in Ch. 4 in connection with deference to experts and conceptual understanding.

⁴⁸ There are some obvious parallels here with the debates between internalists and externalists about epistemic warrant in the case of beliefs. The internalist version of the objection described above can be formulated as a

In the following section I consider an alternative condition that avoids the (probable) hyper-intellectualism of the evidential condition.

2.4.3 A Reliability Condition

I now want to consider an alternative answer to the epistemic challenge based on an objection to the evidential condition raised by Margolis and Laurence (2011). Margolis and Laurence offer what I call “the reliabilist reply.” They write:

“Fodor moves swiftly from the need for a learner’s activities to be rational (in some minimal sense) to the claim that her observations must count as evidence for what she acquires, and from here to the thought that the learner has to register the evidence by computing its bearing on an explicitly entertained hypothesis... Fodor is presupposing an internalist approach to justification. But an externalist approach would suffice just as well as a baseline requirement for explaining the ‘rational’ dimension of concept learning. In that case, what matters isn’t so much that the agent recognizes the evidential value of what she observes, but rather that she employs cognitive mechanisms that deliver new concepts through processes that reliably reflect appropriate environmental contingencies. Whatever one thinks about epistemic justification more generally, there is a good case to be made that when it comes to ‘the rationality’ implicit in concept learning, externalist justification may well be all the justification that’s required.” (2011, p. 519)

Margolis and Laurence suggest that Fodor is presupposing an “internalist approach to justification”. In this chapter I have framed the discussion in terms of the broader category of epistemic warrant. Justification is a species of epistemic warrant. In what follows I will not be concerned with comparing internalist and externalist notions of *justification*. As I stated earlier, I use the term ‘justification’ to refer to a species of epistemic warrant that requires the conceptualization of reasons by an individual (or an individual’s psychology) who employs the justification.⁴⁹ Whether one is inclined to reserve the term ‘justification’ for the species of warrant that I have described is merely a terminological issue. The substantive issue is whether

demand for access to the essentialist assumption or for the essentialist assumption to be represented among the psychological states of the individual acquiring a natural kind concept on the basis of a stereotype.

⁴⁹ This may very well be the notion of justification that Margolis and Laurence are attributing to Fodor. But I doubt that this is the notion of justification they have in mind when they suggest that an “externalist justification may well be all the justification that’s required” (2011, p. 519).

there is a species of epistemic warrant that is both principled and applicable to the case of concept acquisition based on experience. I take it that Margolis and Laurence think there is such a species of epistemic warrant, some version of what they call an “externalist approach” to justification.

An alternative to the evidential condition can be drawn out of a criticism of Fodor put forward by Margolis and Laurence (2011, p. 519). I call this the “concept acquisition reliability condition”:

The Concept Acquisition Reliability Condition (CAR): The acquisition of a concept is warranted if a concept is acquired from cognitive mechanisms using processes that reliably reflect the appropriate environmental contingent relations between kind syndromes and natural kind properties.

The essentialist concept acquisition process does reliably reflect the contingent relations between kind syndromes and natural kind properties. This should not be surprising given that the model was built to exploit the reliable correlation between kind syndromes and natural kind properties. However, CAR by itself could not be a sufficient condition on warranted concept acquisition. If it were sufficient, then the explanation of *how it is* that perceptual experiences mediated the process of acquiring a natural kind concept would be irrelevant.⁵⁰

Consider a version of this worry raised by an objector that I will call the “not-merely-reliabilist.” The not-merely-reliabilist wants some further condition that can distinguish between a process that merely reliably reflects the relations between syndromes and kinds, and processes that do so non-accidentally. To get at the intuition that motivates the not-merely-reliabilist, consider the following case of the fastidious demon:

⁵⁰ To be fair, Margolis and Laurence do not explicitly claim that the reliability condition under consideration is sufficient. Nor do they spell out what kind of reliability condition they have in mind. What I am calling the concept acquisition reliability condition is my reconstruction of the kind of condition to which they would be amenable, given their criticism of Fodor. Though they note the importance of questioning Fodor’s insistence on evidential-inferential relations between stereotypes and acquired concepts, giving an account of the epistemology of concept acquisition is not the focus of their paper.

The fastidious demon: Suppose that there is a fastidious demon watching over us such that whenever anyone comes within 100 feet of an instance of some natural kind the fastidious demon implants the concept that indicates that natural kind, and a representation of the kind syndrome for that kind. The fastidious demon does this for purely aesthetic reasons. She enjoys the elegant symmetry of aligning instances of properties, with concepts and representations of kind syndromes in the minds of nearby cognizers.

In the fastidious demon world, individuals acquire natural kind concepts via a process that reliably reflects the appropriate environmental contingent relations between kind syndromes and natural kind properties. One could object that in this case the concepts are not strictly speaking acquired from *cognitive* mechanisms, because the interventions of the fastidious demon lie outside of our psychologies.⁵¹ I think this objection misses the point illustrated by the example. The fastidious demon is shorthand for some mechanism that delivers a non-luck-based accidental reliable representational connection between natural kind properties, perceptual experiences, and acquired concepts. The twin-factors of the proximity of instances of properties and the predilections of the fastidious demon result in individuals having representational competences that mark abilities for minimal understanding of the concept—understanding sufficient to attribute the concept to the individuals.

The representational connection through representational competences is accidental relative to our representational capacities—that is—accidental relative to processes that aim at veridicality and truth. Veridicality or correctness plays no role in the process by which concepts are acquired in the fastidious demon world. By contrast, the representational connection is not accidental relative to the aesthetic predilections of the fastidious demon. The representations of

⁵¹ Such a mechanism would be what Nicholas Silins calls “incidentally reliable” (2012, p. 252). The fastidious demon case is inspired by Alvin Goldman’s case of the benevolent demon (1976). The benevolent demon makes it so beliefs formed by wishful thinking usually come true. There are some obvious parallels here with the debates between internalists and externalists over warranted beliefs. A weak internalist version of the objection described above can be formulated as a demand for the essentialist assumption to be *represented* among the psychological states of the individual acquiring the concept, though it need not be deployed in a chain of reasoning. This demand need not invoke the strong internalism underlying the evidential condition.

kind syndromes that the fastidious demon provides are not implanted as accurate representations of surrounding kinds. Rather their presence in combination with certain concepts in the proximity of instances of certain kinds generates an aesthetically pleasing tapestry of things in her world. She has arranged them in the way we might arrange a bouquet of flowers—with no regard for accuracy or correctness, but simply in a manner to suit our taste.

The point of the example is that if a condition on forming concepts only requires that the process reliably reflect the relations between kind syndromes and natural kind properties, then that condition will fail to capture the difference between cases of accidental and non-accidental representational competences and connections. As a consequence it allows for reliability that is independent of the representational nature of the psychological capacities that make up the concept forming process, and independent of the constitutive relationship between those capacities and the capacity for forming empirical beliefs. It allows for *mere* reliability as well as reliability derived from exercising distinctively representational capacities. What we want is a condition that only allows for this latter kind of reliability. Only this latter kind of reliability can explain our warrant for what I called early beliefs about the unknown.

Recall that entitlement is warrant by a psychology's operating well with respect to its representational functions. The fastidious demon example shows that CAR by itself fails to exclude explanations that are wholly independent of representational function or representational success.

The thought experiment can, of course, generate competing intuitions. Some will no doubt say that acquiring concepts as a consequence of the fastidious should qualify as a warranted concept acquisition process and thus a process that yields at least *prima facie* warrant. After all, the process leaves us well positioned to pursue empirical knowledge of the world. At

the very least it leaves us well positioned to form true early beliefs about the unknown. Such beliefs could serve as a foundation for forming hypotheses and theorizing about the deeper structure of the empirical world, including the relations between the concepts with which we have been endowed.

Though I can understand the pull of the intuitions underlying the above response to the thought experiment, I think there is room to push back. First, to be clear, I am not claiming that the concept acquisition process by way of the fastidious demon undermines all warrant for forming empirical beliefs. My criticism of CAR does not entail that individuals in the fastidious demon world are doomed to a life of unwarranted empirical belief. For those living in the fastidious demon world, representational success has nothing to do with how they acquire their natural kind concepts, but it can still be part of the explanation of warranted empirical belief formation down the line. For example, even if some early belief such as the belief THAT IS A CAT formed at the first sight of a cat would not have prima facie warrant on my view, with time and more cat experiences (and experiences with others who have engaged with cats) individuals in the fastidious demon world could come to form many warranted beliefs about cats.

One could object that the preceding paragraph amounts to a concession that renders the prima facie warrant I have been describing superfluous to an explanation of warranted empirical belief. For now I will make just two points in response to this charge. I will return to the objection in section 6 after I present my account of concept formation entitlement.

First, suppose that the concept acquisition process in the fastidious demon world does qualify as an epistemically good process by which to acquire concepts. But in our world, given that we can produce evidence that there is no active fastidious demon, there remains the question of how we can acquire or form concepts in a way that yields the prima facie epistemic warrant

for employing natural kind concepts in (at least) our early beliefs about the unknown.⁵² So even if early beliefs about the unknown were deemed warranted in the fastidious demon world, that would not resolve the question with respect to our world.

My second point is more general and concerns the scope of epistemology. If epistemology is tasked with the job of mapping all the components and cases of warranted belief, then early empirical beliefs about the unknown provide a case which epistemology must account for. I am offering an account of the source of the warrant for such beliefs. The fact that a failure to have warrant for such beliefs does not undermine the possibility of ultimately forming warranted empirical beliefs does not show that early beliefs are of no interest to epistemology. In so far as epistemology seeks to identify and explain sources of warrant, an explanation of the source of the warrant of early beliefs is needed.

I am arguing that an explanation of that prima facie warrant must be grounded in the notion of representational success. The reliability achieved in the case of the fastidious demon is devoid of any sense of representational achievement. Individuals living in the fastidious demon world might acquire the concept CAT as a consequence of being in the proximity of an instance of the property, but successful representation of the objective empirical world, as a function or as a goal, has no role to play in the explanation of why they do so.

If the reliability of a concept acquisition process is to be part of the answer to the epistemic challenge, then that reliability must be grounded in the operations and function of a psychological process. Otherwise the reliability will be accidental relative to our representational

⁵² We can produce evidence to the extent that we can ever produce evidence that an individual have some particular concept C. To show that there is no active fastidious demon in our world we just need to show that there is an individual that fails to have some natural kind concept that indicates a property that has been instantiated at some time within the 100-foot radius of the individual in question.

capacities. What we need is an account of the source of the normative dimension of a warranted psychological process.

2.4.4 Representational function and reliability

In this section I present Tyler Burge's account of perceptual entitlement. I do this for two reasons. First, Burge's account of perceptual entitlement illuminates the shortcomings of the reliability condition. And crucially, my account of concept formation entitlement is built on Burge perceptual entitlement. This section will proceed as follows: I present a short overview of Burge's project and his background teleological framework for representational capacities. I then explain how Burge's framework can make a principled distinction between accidental and non-accidental reliability. I end by returning to the example of the fastidious demon.

Burge's account of perceptual entitlement is an account of the warrant of the formation of *perceptual beliefs*. The epistemic norms that Burge describes are the norms governing the transitions from perceptual states to the formation of perceptual beliefs. The type of conceptualization that Burge discusses is a *strict conceptualization* of perceptual attributives that indicate perceptually discriminable properties such as shape properties, color properties, textural properties, locations, and motion. The result is a basic perceptual belief, one that does not go beyond the content of the perceptual state in any way beyond the form of the state (2003, pp. 541-5).⁵³ A basic belief is a conceptualization of a perceptual state. Conceptualization is a process by which perceptual representation is converted into propositional representation. Thus, a basic perceptual belief can only attribute properties that are represented in perception such as

⁵³ The distinction between a perceptual representational content and its conceptualized counterpart is crucial for taxonomizing the various relevant psychological kinds that figure in psychological explanation. Conceptual content is able to figure in more complex kinds of states and transitions than perceptual content. The scope of this paper prevents me from giving a lengthy defense of this claim. For detailed discussions of Burge's distinction between perceptual and conceptual states see Burge (2010a, 2010b).

motion pattern, shape, color, texture, location, and the property of being a body. The account of conceptualization of the perceptually discriminable properties he describes suggests only the beginnings of an explanation of the warrant for the acquisition of natural kind concepts. Since I am not claiming that there are perceptual attributives that indicate natural kinds, the transitions that are part of natural kind concept formation are not instances of Burge's strict conceptualization.⁵⁴

Burge's framework grounds norms in the representational function of some psychological capacity and the anti-individualistic nature of our perceptual states and perceptual beliefs.

Anti-individualism in its most general form states that:

“The natures of many mental states constitutively depend on relations between a subject matter beyond the individual and the individual that has the mental states, where relevant relations help determine specific natures of those states.” (Burge, 2010a, p. 61)

However, because I have restricted the scope of this chapter to empirically derived natural kind concepts and their role in empirical belief formation, the specific form of anti-individualism that is relevant for our discussion states that:

“The natures of mental states that empirically represent the physical environment depend constitutively on relations between specific aspects of the environment and the individual, including causal relations, which are not in themselves representational; the relevant environment–individual relations help determine specific natures of the states.” (Burge, 2010a, p. 61)⁵⁵

According to Burge, warrant, (and thus entitlement), “are fulfillments of norms associated with achieving the representational good, truth” (2003, p. 506). Burge offers a teleological account of representational states that contribute to epistemic entitlement to hold perceptual beliefs on the basis of perceptual experience. Burge draws a connection between the

⁵⁴ Burge does not address the question of whether strict conceptualization is a process by which concepts that correspond to perceptual attributives are *acquired*.

⁵⁵ For a detailed discussion of the view see Burge (2010a, Ch. 3).

epistemic norms that govern representational states and processes, and the representational function of such states and processes to represent veridically. As he explains,

“[Epistemic norms] are standards for the formation of certain representational states, processes, and methods in fulfilling the representational function of representing veridically. They are standards for representing reliably and well, given the capacities and perspective of the individual or subsystem. ...[Epistemic norms] do not rest purely on reliability. They apply when and only when certain psychological competencies are present. Given the representational abilities of the animal, and given its perspectival limitations, there are standards that govern states and transitions that well serve the function of forming true beliefs.” (2003, p. 513)

Burge is interested in the contribution that perceptual states make to the epistemic entitlement to hold perceptual beliefs. He argues for perceptual *entitlement* (as opposed to justification) because he rejects the view that perceptual states provide reasons for perceptual beliefs. According to Burge there are two conditions on perceptual states that explain their distinct contribution to the epistemic entitlement of perceptual beliefs: an *individuation condition* and a *reliability condition*.

Burge’s anti-individualism informs his individuation condition. He gives the following two-part condition on perceptual states having the content that they have. There must have been,

[i] “causal-formative interactions (which are not in themselves representational)” and

[ii] “Representationally successful interactions between instances of types of relevant perceptual referents and aspects of the individual’s perceptual system (in either the individual’s history, or in the evolution of the system in his evolutionary ancestors, or in some other way).” (2003, p. 531)

According to Burge’s *reliability condition*, perceptual states have to be reliably veridical in a creature (or subsystem’s) normal environment in order to contribute to entitlement. The normal environment is the environment by reference to which the representational content of a perceptual state “is explained and established” (2003, p. 536). Burge argues that reliability in some other environment or under some other conditions could not contribute to entitlement

because such reliability would be accidental relative to the nature of the perceptual states. Burge explains:

“The idea behind this view of entitlement is that reliance on perceptual states is warranted partly because the very identity of the states is constitutively and explanatorily associated with veridical representation and with guiding activity with respect to entities veridically represented... Veridicality enters into the very nature of perceptual states and abilities. So it is built into the nature of the competence associated with the formation of a reliable perceptual state that the state make a non-accidental, explanatory, positive contribution to true belief and knowledge in the animal's normal environment in animals that are capable of true belief and knowledge.” (2003, p. 532)

Burge's account of perceptual entitlement provides a principled way to distinguish between merely reliable representational states and reliable representational states that contribute to epistemic entitlement. States of the latter sort are constitutively and explanatorily associated with a representational system's function to represent veridically in the environment in which the representational content of the reliable state is explained and established. Whether some type of state is in fact reliable is *metaphysically contingent*. But if the state has a propensity to be reliable in normal conditions, its reliability is non-accidentally related to the state's nature.

Burge's account of the relationship between the individuation condition and the reliability condition is instructive when comparing notions of reliability. Recall the formulation of the reliability condition (CAR) that I attributed to Margolis and Laurence:

The Concept Acquisition Reliability Condition (CAR): The acquisition of a concept is warranted if a concept is acquired from cognitive mechanisms using processes that reliably reflect the appropriate environmental contingent relations between kind syndromes and natural kind properties.

CAR does not connect reliability with the individuation of the concepts acquired. In their defense, recall that Margolis and Laurence are criticizing Fodor and thus assuming as background as much of Fodor's own view as possible in levying their criticism. Part of that

background includes Fodor's information-based theory of content.⁵⁶ Information-based atomism cannot give the reliability condition the kind of support that Burge's anti-individualist individuation condition provides. As I noted earlier in section 3, Fodor's information-based atomism's the explanation of why experiences of some property p leads to the acquisition of the concept that indicates p makes no mention of representational success. That is, on Fodor's view, that the experience of cats leads to the acquisition of the concept CAT is arbitrary with respect to successful representation of the empirical world.⁵⁷

Without supplement, CAR cannot exclude explanations that are wholly independent of representational function or representational success. For example, the reliability achieved in the case of the fastidious demon does not reflect representational success. As I remarked earlier, my criticism of CAR does not entail that individuals in the fastidious demon world are blocked forever from the possibility of warranted empirical belief. But it does entail that the process by which they acquire their natural kind concepts cannot be a source of warrant for their early empirical beliefs that employ their newly acquired natural kind concepts

2.5 Concept formation entitlement

In this section I present my account of what I call "concept formation entitlement." My account of concept formation entitlement is modeled in part after Burge's account of perceptual entitlement. I assume as background Burge's teleological framework and the explanatory relationship between reliability and the nature of representational states understood anti-individualistically. There are important differences between Burge's account of perceptual

⁵⁶ Laurence and Margolis (2002) argue that "theories of content and theories of concept acquisition need to be juxtaposed" and that much of the debate surrounding concept acquisition has failed to do this (p. 28).

⁵⁷ Or as Fodor (1998) suggests, it is grounds for thinking that properties such as the property of being a cat are mind-dependent. As I noted in section 3 footnote 19, I take Fodor's argument for mind dependence to be a non-starter. Some comfort can be taken in the fact that in more recent work Fodor seems to retreat from this argument. (See for example, Fodor and Pylyshyn 2014).

entitlement and my account of concept formation entitlement. The differences stem mostly from the differences in the representational kinds and transitions to which our respective accounts apply. Still I take my account to be an extension of Burge's account of perceptual entitlement.

What my account has in common with Burge's account is a framework that situates some psychological process in a broader teleological understanding of psychological capacities. I claim that the framework that Burge describes for the conceptualization of perceptual attributives and the formation of perceptual beliefs can be extended to cover other psychological processes where the nature of the process is constitutively associated with the representational function of representing veridically. Furthermore, because I am interested in the warrant for concept acquisition on the basis of perceptual experience, Burge's account of the contribution that perceptual states make to the warrant of perceptual beliefs is foundational to my project. I developed the notion of concept formation entitlement in part to remedy the shortcomings of the concept acquisition reliability condition.

I will now detail my positive proposal for a notion of warrant that I call concept formation entitlement. Concept formation entitlement (CFE) is a kind of warrant that can attach to concepts formed on the basis of perceptual experience including perceptually derived representations such as stereotypes. CFE is an externalist notion of warrant in that part of what makes an individual warranted includes relations between that individual and the environment.

On my proposal, the entitlement (and thus warrant) achieved in forming natural kind concepts on the basis of perceptually derived representations has two sources: (1) the nature and function of our sense-perceptual capacities by which perceptual states and stereotypes are formed; and (2), the contribution that the process makes to the fulfillment of epistemic norms

that govern the formation of empirical beliefs and empirical discovery. I will discuss each of these sources briefly.

2.5.1 Source1: The nature and function of our sense-perceptual capacities

The perceptual basis of concept formation mediates a connection to the empirical world. Stereotypes are representations that are formed on the basis of basic perceptual beliefs. Basic perceptual beliefs are formed on the basis of perceptual states. Thus perceptually derived stereotypes are causally connected to the world through the exercise of representational competences of the perceptual system. Stereotypes, in turn, causally connect concepts to the world.

The representational function of the perceptual system is to represent veridically and thus reliably. Crucially, the perceptual system does more than provide a causal connection to empirical subject matter. Namely, the individuation and objectification of empirical subject matter begins in perception. An experience as of a body with some particular shape, motion pattern, coloring, and texture purports to be tracking some objective thing in the world. This is the first step in positing an objective subject matter for empirical investigation.

The essentialist concept formation process depends on the perceptual system's being generally reliable at least with respect to instances of the salient observable features that figure in stereotypes and in the positing of particular bodies. If the perceptual system were generally not reliable with respect to these features, then there would be no reason to expect that stereotypes would reliably reflect the kind syndromes out in the objective empirical world. An experience as of a body with some particular shape, motion pattern, coloring, and texture purports to be tracking something in the world. The process presupposes that where you have frequently co-instantiated features it is a good bet to posit a kind. (Of course whether there is actually a kind (or property) that can be reliably tracked using some particular set of frequently co-instantiated

features is ultimately an empirical question). Thus in addition to generally reliable representation of objective particulars, the process presupposes veridical and reliable representation of the observable features that are represented in stereotypes. So concept formation entitlement is underwritten by our entitlement to rely on perceptual states generally, and perceptually derived representations such as stereotypes.

The point of the preceding discussion of representational functions and epistemic norms was to set up a framework that shows that forming natural kind concepts on the basis of perceptual experience reflects a kind of epistemic doing well, and that the explanation of the epistemic doing well is not merely a matter of being reliable. The perceptual basis of natural kind concept formation is at least part of the explanation of the non-accidental reliability of empirical discovery. Thus the perceptual basis of natural kind concept formation plays a fundamental role in concept formation entitlement.

The claim that natural kind concepts are formed on the basis of perceptual experience is not sufficient to underwrite a principled distinction between accidental and non-accidental reliable processes that is needed to answer the epistemic challenge. The principled distinction between accidental and non-accidental reliability requires that perception—as a representational capacity—is understood anti-individualistically. To see why, consider why Fodor cannot make the same principled distinction just from the claim that natural kind concept acquisition is *caused* by the formation of stereotypes *formed on the basis* of perceptual experience. The distinction between accidental and non-accidental reliability turns on Burge's individuation condition. There must be a privileged environment with respect to which representational content is explained and established. What Fodor needs is a causal connection that is non-accidental relative to the representational content of the concept acquired. But Fodor's information-based atomism does

not have a privileged environment because the manner in which concepts are acquired on his view is independent of any particular environment. So on his view perception does not bear the appropriate relation to the environment to contribute to and underwrite concept formation entitlement for natural kind concepts.

The distinction between accidental and non-accidental reliability turns on more than just Burge's individuation condition. Burge's framework is not merely a causal-historical view of the individuation of mental states. It is a constitutive account—the full epistemological upshot of which cannot be understood independently of the teleological account of the psychological capacities for the representation of empirical objective subject matter, and empirical belief formation.

2.5.2 Source2: Making a contribution to the epistemic norms of empirical discovery

I said that concept formation entitlement has two sources. The second source is the contribution that a concept formation process makes to the fulfillment of epistemic norms that govern the formation of empirical beliefs and empirical discovery. As part of the epistemic endeavor of discovery, natural kind concept formation must produce concepts that reflect the observer-independent nature of the properties that they indicate, and do so in a way that respects the inherent epistemic distance between a representation of an objective empirical subject matter, and the nature of that objective empirical subject matter. The nature of an objective particular or an objective attribute is independent of how it is represented, how it appears, how it tends to behave, etc. The objective empirical world is the final arbiter of the nature of empirical subject matter. In this section I will suggest one such set of norms that I call “revisability norms”.

I said that we should think of the aim of discovery as learning the truths of the empirical world, and in particular the truths that are needed for understanding the nature and structure of the empirical world. That aim does not preclude getting things wrong en route. Empirical beliefs

are by their nature fallible and defeasible. Hypotheses are meant to be tested, leaving open the possibility that they will be rejected. Natural kind concepts must be formed in a way that minimizes any commitments that could undermine the possibility of hypotheses disconfirmation or the revisability of the empirical belief in which a natural kind concept figures.⁵⁸ For example, if the belief that FS ARE GS is revisable, then there must be an epistemically warranted route to rejection, given some possible empirical evidence. Arithmetic beliefs provide a helpful contrasting example. A belief such as TWO PLUS TWO EQUALS FOUR is not open to revision such that there must be an epistemically possible warranted route to its negation given some empirical evidence.⁵⁹

To contribute to the fulfillment of the epistemic norms that govern empirical discovery, a cognitive system must form concepts that carry minimal non-defeasible ontological commitments about the empirical world. For any natural kind concept C that indicates a property p, fewer ontological commitments packed into C increases the potential revisability of beliefs about p. For example, in order to be able to discover that ripe oranges can be green (if grown in tropical climates), the belief that ALL RIPE ORANGES_{NK} ARE ORANGE_{COLOR} must be revisable. It would not be revisable if the natural kind concept ORANGE_{NK} carried any non-defeasible ontological commitments about the color of oranges.⁶⁰

⁵⁸ One could characterize this as minimizing the analytic truths that are entailed by natural kind concepts. Strevens 2012 develops an account of how a concept could be acquired without the creation of analytic truths. I developed my view independently of, and prior to learning about Strevens' view. Strevens' view is entailed by my view, but my project includes the stronger claim: that such concept acquisition processes are epistemically evaluable and can make a positive epistemic contribution to empirical beliefs.

⁵⁹ There need not be an *actual* rational route to revising the belief that FS ARE GS. It could be the case that actually all Fs are Gs, and no one in the actual world has any reason to believe otherwise.

⁶⁰ As I mentioned previously, there is no shortage of examples where a subject S is grossly ignorant of true facts about some property such as the property of being a cat, and yet it seems appropriate to ascribe to S mental states with the representational content CAT. For the original examples see Putnam (1962/1975a), (1962/1975a), (1973/1975b), (1975/1975d) and Burge (1979/2007a), (1982/2007c), (1986/2007b). I discuss the limits of revisability of empirical beliefs in greater detail Chapter 3.

The essentialist concept formation process produces natural kind concepts that contribute to the fulfillment of revisability norms. Part of the explanation of why this is so is that the process preserves the distinction between representations of syndromes or purported syndromes of a property and the property itself. None of the superficial features that are represented in the stereotype are treated as constitutive of the content of the natural kind concept acquired on the basis of that stereotype. So the process succeeds in minimizing the ontological commitments of empirical beliefs and in preserving the empirical discovery system's sensitivity to empirical evidence. Thus the process can produce concepts that make a positive epistemic contribution to fulfilling epistemic norms of revisability.

So, for example, natural kind grouping privileges the transition from the cat stereotype to the natural kind concept CAT by treating the statistically co-instantiated salient features represented in the cat stereotype as merely symptomatic of a naturally occurring discoverable property, rather than treating these features as defining or constructing a category. A grouping that treated the features represented in the cat stereotype as criterial for group membership would yield the SHMAT concept described above.

2.6 An objection to the need for concept formation entitlement

Before concluding I want to consider the objection that I noted at the end of section 4. The objection challenges the need and thus legitimacy of incorporating an epistemology of concept formation into the broader project of the epistemology of empirical knowledge. I believe that this objection underlies one of Jerry Fodor's few meditations on the connection between the process by which a concept is acquired, and the use of a concept in empirical investigation.

Fodor considers the following question:

“What reason do you have for supposing that our concepts apply—that our science is true—if warranted inference doesn't mediate their attainment?”

(Part of) Fodor's answer is the following:

“[T]he truth of our science doesn't depend on how we came to discover it, only on its correspondence to the facts. The origin of a concept is, no doubt, a poor predictor of its warrant; so you should establish warrant by argument to the best explanation, not by committing the genetic fallacy. A Rationalist would predict that the history of theory construction, and of the ontogeny of our concepts, might well be a matter of sudden insights and flashes of intuition; of inspiration... [But this] affects the truth of our science not one whit, so long as the traditional distinction between the logic of discovery and the logic of confirmation is scrupulously observed.”(1981, pp. 313-4)

Fodor is correct that the truth of our science depends “only on its correspondence to the facts.”

But the truth of our science is neither sufficient nor necessary for the warrant of our science.

Standards for truth are distinct from standards for warrant. The passage above comes from Fodor's reflections on the demise of the empiricist dream of building concepts out of sensory primitives by logical construction. But the spirit of Fodor's comments translates easily into an objection to any theory that claims that the origin of a concept bears on our prima facie warrant for forming empirical beliefs that employ newly formed natural kind concepts.

At the beginning of this chapter I motivated the project of developing an epistemology of concept formation by arguing that we need an explanation of the source of our warrant for forming early beliefs about the unknown. Fodor would likely say that early beliefs are just weakly held hypotheses whose warrant will bear out if confirmation levels are reached. Fodor recommends that we scrupulously observe “the traditional distinction between the logic of discovery and the logic of confirmation.” The advice paints a picture in which explaining the origin of the newly formed concepts that we employ in forming our early beliefs about the empirical world is a distant explanatory project, unconnected to our explanation of the warrant of such beliefs. On this picture our representational abilities are like tiles in a magnetic poetry board. We combine the tiles into propositions and our warrant for believing any of those propositions turns on our confirmation practices. This leaves a disconnect between our capacity

for forming the tools we need to investigate the world and our warrant for holding the products of that investigation.

In my discussion of the fastidious demon world I said that those living in the fastidious demon world were not doomed to a life in which warranted empirical belief was impossible. Lack of prima facie warrant to employ newly acquired concepts in the formation of empirical beliefs—even early empirical beliefs—is not an epistemic roadblock to eventual warranted empirical belief and thus the possibility of scientific knowledge.

I noted that this invites the objection that the prima facie warrant I have been describing is superfluous to an explanation of warranted empirical belief. I think the objection misses the significance of the epistemic contribution that concept formation entitlement makes when concepts are formed in epistemically good ways. Earlier I characterized that contribution as the epistemic bridge between perceptual beliefs, and our ability to form empirical beliefs that go beyond what we can represent in perception. Just as we need an explanation of the warrant of perceptual beliefs, we need an explanation of our warrant for going beyond perceptual beliefs. Though a lack of concept formation entitlement does not always undermine warrant for empirical beliefs, it does play an explanatory role in our understanding of the epistemic achievement of warranted empirical belief. That achievement is constitutively bound up with the representational function of our representational capacities, including our capacity for forming new ways of representing.

2.7 Conclusion

I have been arguing that the epistemology of concept formation is part of the epistemology of empirical investigation. Part of what it is to discover the world and its deeper structure is to form new ways of representing the world and its structure. I claimed that we need an epistemology of concept formation to explain the epistemic continuity from perceiving the

world to achieving empirical knowledge. That there could be good or bad ways (epistemically speaking) of forming concepts does not seem so strange when looking at the pursuit of empirical knowledge from the ground up. In the non-accidental case, forming a concept through the use of representational capacities to represent objective particulars and attributes reflects the connection between our ability to represent objective properties, and our ability to represent particulars that instantiate such properties. Merely having a concept does not reflect that connection. Getting a representational competence in the right way—the epistemically good way that comes with *prima facie* warrant to employ that concept—requires that the formation process be non-accidental relative to our capacities for representing objective empirical subject matter.

In arguing that the process of concept acquisition is epistemically evaluable, I am putting forward a view on which the epistemology of subpropositional representational elements is ultimately part of an explanation of the possibility of warranted empirical belief and knowledge. Knowledge of the empirical world begins with veridical representation of objective subject matter. Epistemic norms do not just appear at the level of beliefs disconnected from the nature of our representational capacities. Concept formation entitlement is part of the explanation of the positive contribution that sub-propositional representational components make towards the warrant of an empirical belief. While the process by which a concept is formed is not part of the explanation of the *truth* of a true empirical belief, it is part of the explanation of the warrant of an empirical belief and thus ultimately part of the explanation of scientific knowledge.

Chapter 3

Concept Formation Entitlement and Belief Revision

Introduction

In Chapter 2 I argued that natural kind concept formation is part of the greater epistemic endeavor I called “empirical discovery”. Empirical discovery includes concept formation, belief formation, belief revision, conceptual mastery, and theory building. The goal of empirical discovery is warranted empirical belief and ultimately, to achieve an understanding and knowledge of the empirical world. I said that part of what it is to discover the world is to form new ways of representing the world and its structure. And I argued that an explanation of empirical discovery has to begin with an explanation of how the representational system picks out particulars in perception, and creates ways of representing those particulars as members of a kind.

I claimed that we need an epistemology of concept formation to explain the epistemic continuity from perceiving the world to achieving empirical knowledge. Achieving this epistemic continuity means filling in a glaring gap in our epistemology of the pursuit of empirical knowledge. That gap is the transition from basic perceptual beliefs to empirical beliefs that require the formation of new concepts. Basic perceptual beliefs are beliefs whose representational content indicates only perceptually discriminable properties, such as the property of being round or the property of being a body. Perceptual entitlement entitles the individual (or her psychology) to hold the basic perceptual belief based on a perceptual state. Our empirical beliefs go far beyond beliefs about such perceptible properties. In order to transition from basic perceptual beliefs to beliefs about properties that are not represented in perception we must form new concepts. Call the first beliefs that are formed using a new concept “early

beliefs”.⁶¹ An epistemology of that transition from basic perceptual beliefs to early beliefs is the missing epistemic link in an epistemology of empirical discovery and ultimately knowledge. Concept formation entitlement is the prima facie warrant to form early beliefs using newly formed concepts.

In Chapter 2 my proposal was restricted to natural kind concepts. In this chapter I present my account of concept formation entitlement in more detail and I expand my proposal to cover a larger class of concepts, what I call “GEM concepts” (or GEMs for short). As I explain in more detail below, GEM concepts are general concepts that indicate objective properties that are instantiated in whole particulars that can be picked out in perception. Because the class of GEM concepts includes many concepts that are not natural kind concepts, Margolis’s essentialist model that I described in Chapter 2 will have to be replaced.

The rest of the chapter will proceed as follows: In part I I explicate the notion of a GEM concept and I explain the foundational role of perception in the GEM concept formation process. In part II I present a model of GEM concept formation. The model is informed by research in perceptual, developmental, and cognitive psychology. In part III I re-introduce the notion of concept formation entitlement and explain how a concept formation process can epistemically support initial concept application in the formation of our earliest empirical beliefs. Concept formation entitlement in part explains how we can be warranted in forming such beliefs in the face of tremendous ignorance. Finally in part IV I consider some objections and offer my replies.

⁶¹ Here I have in mind early beliefs formed by individuals with limited representational resources such as infants, toddlers, and young children. Though a mature adult will have some of these early beliefs, they undoubtedly could form sophisticated complex beliefs by relying on a large stock of background knowledge and an extensive repertoire of concepts.

3.1 Part I: What are GEM concepts?

3.1.1 General, empirical, mediumish

GEMs are general concepts that indicate objective properties that are instantiated in particulars that can be demonstrated in perception. GEM concepts indicate a variety of properties including: natural kind properties such as [^]cat^{^62}, [^]lemon[^], and [^]koala[^]; artifact kind properties such as [^]carabineer[^], [^]knish[^] and [^]phone[^]; as well as an assortment of other properties that defy easy classification such as [^]mountain[^], [^]rock[^], [^]lake[^], and [^]puddle[^]. The property indicated by a GEM concept is paradigmatically instantiated in a whole physical particular. Whole physical particulars are individuated relative to perceptual capacities. The acronym ‘GEM’ stands for ‘general,’ ‘empirical,’ and ‘mediumish’. I employ the acronym to remind the reader of the three signature characteristics of GEMs. These signature characteristics hold in the paradigmatic instances of GEM concepts. I explicate the notion of a GEM concept by discussing each of these characteristics in turn.⁶³

GEM concepts are *general* in at least the following three senses: First, GEMs are predicative concepts that function to indicate properties or kinds (as opposed to singular concepts that indicate particulars); and to attribute the properties they indicate. Second, GEMs are general in that their form and representational content allows them to apply to indefinitely many particulars. Third, GEMs are general in the sense that in forming a GEM concept, a representational competence is formed, which is exercisable (at least in principle) repeatedly.⁶⁴

⁶² I use ‘[^]x[^]’ as a shorthand for the cumbersome locution “the property of being an x.” I use the convention of small caps to mention concepts.

⁶³ At this point in the paper it will be helpful to avoid an extended discussion of non-paradigmatic cases.

⁶⁴ Here I am employing three notions of generality distinguished and explicated by Tyler Burge (2009). Burge calls these notions of generality syntactic generality, semantic generality, and ability generality (respectively). For a detailed discussion of these notions see Burge (2009, pp. 258-72).

GEMs are *empirical* concepts in the following two senses: (1) they indicate properties that are the subject matter of empirical investigation; and (2), they are constitutively associated with the formation of *empirical beliefs*. Empirical beliefs are characterized by the source of their potential warrant. (When warranted) the warrant for empirical beliefs is at least in part dependent for its force on sense experience. Nearly all empirical beliefs have the objective world as their subject matter. On my usage the belief that, IF CATS ARE FURRY, THEN CATS ARE FURRY, does not qualify as an empirical belief despite having some part of the objective world as part of its subject matter because the warrant for this belief is not dependent for its force on sense experience. Some examples of empirical beliefs are CATS ARE FOUR LEGGED, MOUNTAINS ARE FORMED BY VARIOUS KINDS OF TECTONIC ACTIVITY, and KOALAS SLEEP IN TREES.

The particulars that instantiate properties indicated by GEMs are *mediumish* in size. I use ‘mediumish’ as a technical albeit somewhat imprecise notion inspired by Quine (1960). The class of mediumish objects is likely bigger than Quine’s middle-sized middle-distance objects. Most mediumish objects are somewhere between the size of a flea and the size of a mountain range (inclusive).

Being mediumish in size is understood relative to perceptual abilities. This condition is not *ad hoc*, given the role of perception in the GEM concept formation process. Mediumish objects are perceptually discriminable independently of theory (though not necessarily independently of some instrument). For example, cells as seen under a microscope could count as mediumish, but a cell is not a paradigmatic instance of a mediumish particular. The concepts STAR and CELESTIAL BODY are also non-paradigmatic GEM concepts. The properties ^{^star^} and ^{^celestial body^} are instantiated in whole particulars that are much larger than mountain ranges. However, because such particulars can be demonstrated in perception (either with or without an

instrument), STAR is a GEM concept. By contrast, electrons and quarks are not mediumish because they cannot be demonstrated in perception even with the aid of an instrument. So ELECTRON and QUARK are not GEM concepts.⁶⁵

3.1.2 Kind sortal GEM concepts

In order to simplify the subsequent discussion and presentation of a GEM concept formation model, I will limit the discussion to a subset of GEM concepts. This subset of GEM concepts are often called “sortal concepts.” Sortal concepts are the concepts that one needs in order to answer questions like “What is it?” “How many are there?” and “Is it the same?” Paradigmatic GEM concepts are kind sortal concepts. They are kind concepts in that they correspond to categories that are inductively richer than concepts that pick out properties such as ^a red thing^ or ^a squishy thing^.

I will restrict the subset further to GEM concepts that correspond to count nouns in English such as ‘cat’ or ‘knish’. Thus I will not be discussing concepts that correspond to mass nouns in English such as ‘water’, ‘gold’, or ‘sand’.⁶⁶ Count nouns can be pluralized and are used to pick out discrete countable units. Mass nouns require some kind of modifier to be counted. For example, you cannot count the coffee on the table but you can count the cups of coffee on the table.

Finally, I will focus the discussion on kind sortal GEM concepts that indicate what are often called “enduring” properties such as ^dog^ as opposed to “transient” properties such as ^puppy^. (Carey, 2009, pp. 263-4) So the subset of GEM concepts I will focus on are kind

⁶⁵ The above is meant as a preliminary introduction to the notion of GEM concepts. Once the GEM concept formation model is detailed the notion of a GEM concept will become clearer.

⁶⁶ Interestingly, not all English count nouns are count nouns in other languages. There are, however, some terms that are mass terms in every language that marks the count/mass distinction. These mass terms pick out materials that are non-solid and non-cohesive such as liquids, powders, and gels. (Huntley-Fenner, Carey, and Solimando, 2002, p. 205)

sortals that indicate enduring properties. However, for ease of exposition I will just use the expression ‘GEM concept’ throughout the paper.

Though the concept formation model I will describe should apply to all GEM concepts, I focus my discussion on GEM concepts that are formed in early development. Ontogenetically early concepts are of particular interest because it is our earliest formed concepts that most strikingly exemplify the mental feat of forming a new way of representing in the face of massive ignorance.

3.1.3 The role of perception in GEM concept formation

I said that being mediumish in size is understood relative to perceptual abilities. Mediumish particulars are individuated in perception. A mediumish particular is demonstrable in perception and can be represented *as* a body. In much of the developmental literature, it is thought that the attribution of bodyhood to particulars is constrained by properties such as cohesion and boundedness⁶⁷. There is some debate over whether the attributive is perceptual or conceptual. It is not essential to my proposal that representation as of a body be available in perception. For present purposes I will assume that it is.⁶⁸

Perception’s role in empirical discovery goes beyond individuating mediumish particulars. Empirical discovery begins with the objective representation of the empirical world. The formation of perceptual states and basic perceptual beliefs is where objective empirical representation begins.

⁶⁷ For a discussion of some experiments that distinguish between non-cohesiveness and non-rigidity see Huntley-Fenner, Carey, and Solimando (2002).

⁶⁸ The assumption is informed by Tyler Burge’s explication of the perceptual attributive *body*. Burge mounts a detailed defense of the view that *body* is a perceptual attributive (see in particular Burge 2010, pp. 454-64 and 2011, p. 125). For arguments in favor of a conceptualist account of the attributive see Spelke 1996. See also Scholl 2001 and Carey and Xu 2001 for a discussion of these issues. For a discussion of some experiments that distinguish between non-cohesiveness and non-rigidity see Huntley-Fenner, Carey, and Solimando (2002).

3.2 Part II: How to Form a GEM Concept, the empirical model⁶⁹

I will now propose a model that captures an empirically plausible account of GEM concept formation. I say that the model “captures” as opposed to describes an empirically plausible account because the proposal is presented in a far more general form than models found in the developmental literature. I will attempt to give the contours of a model that could be fleshed out in a nativist or empiricist direction.⁷⁰ For the purposes of my project I just need a sketch of an empirically plausible model with enough detail to facilitate the discussion of concept formation entitlement in Part III.

The perceptual and linguistic cues and biases described in the model derive from research in perceptual, developmental, and cognitive psychology. After introducing the model I review some of the empirical evidence that points to the availability of the psychological competencies described in the model.

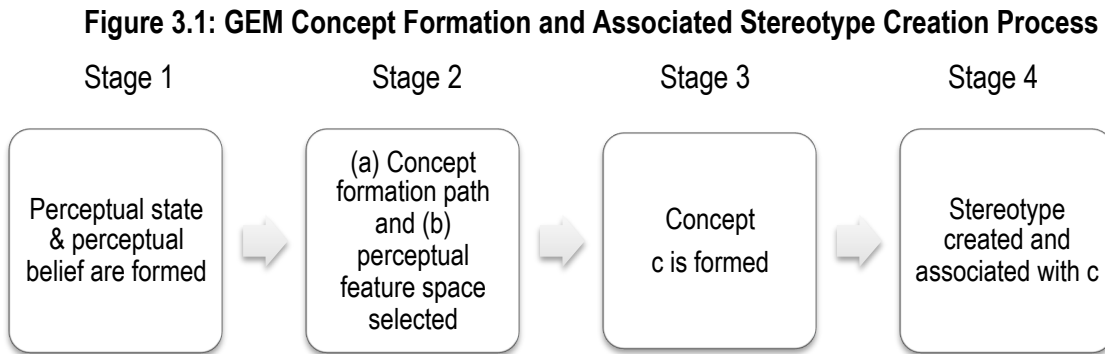
3.2.1 A model of GEM concept formation and stereotype creation

I will begin with a brief overview of the model’s four stages. Then I will go back and discuss each stage in more detail. The model is a model for GEM concept formation *and* stereotype creation. It begins (stage 1) with the formation of a perceptual state and a restricted perceptual belief. At stage 2, one of three types of global concept formation paths is selected (2a). Each of the global paths has an associated perceptual feature space in which salient perceptual features from the representational content of the antecedent perceptual state are

⁶⁹ As I noted in Chapter 2, I do not assume that *all* cases of GEM concept acquisition are dependent on direct experience of a particular that instantiates the property indicated by the acquired concept. For example, the concept OCTOPUS can be acquired by being told a story about an octopus or by being shown pictures of an octopus. My account of concept formation entitlement also applies to cases in which GEM concepts are formed in these other ways. I believe that the explanations of these other ways are ultimately parasitic on the explanation of the direct experience case.

⁷⁰ The model I will describe should be general enough so that it is compatible with both Mandler’s empiricist account (2004) and Carey’s nativist account (2009) of concept formation.

recorded (2b). In stage 3, some new concept C is formed. After C is formed, a stereotype of the extension of C can be created and associated with C . Thus the coarse-grained characterization of the process is as follows:



I will now explain each of these stages in detail and describe some evidence that supports the empirical plausibility of the model.

3.2.1.1 Stage 1: Formation of a perceptual state and a basic perceptual belief

The perceptual system contributes to the concept formation process by forming the first objective representation as of a particular and its perceivable attributes. GEM concept formation begins with the formation of a perceptual state with representational content that indicates a particular and a basic perceptual belief based on the perceptual state. A basic perceptual belief is a belief whose representational content does not go beyond the representational content of a perceptual state in any other way but the form of the state itself. A basic belief is a conceptualization of the perceptual state. Conceptualization is a process by which perceptual representation is converted into propositional representation. Thus, a basic perceptual belief can only attribute properties that are represented in perception such as motion pattern, shape, color, texture, location, and the property of being a body (described earlier). For example, suppose you are walking in the woods and come across a small spherical orange fruit that is in fact a kumquat. You form a perceptual state (PS_k) with the content ***That spherical, orange, body, there***.

(PS_k) is the perceptual state that serves as the basis for the basic perceptual belief (PB_k) THAT IS A SPHERICAL, ORANGE, BODY, THERE. The representational content of (PB_k) introduces no new attributives and differs from the representational content of (PS_k) only with respect its propositional structure. (PB_k) does not attribute a property that is not represented in (PS_k). Suppose that you identify the fruit as a kumquat on the basis of its appearance. You then form the belief (B_k) THAT IS A KUMQUAT. (B_k) is a belief that goes beyond the representational content of (PS_k) in more than just form. Specifically, (B_k) attributes a property that is not represented in (PS_k), the property of being a kumquat.⁷¹

3.2.1.2 Stage 2a: Selecting a concept formation path

The representational content of the perceptual state (and/or the basic perceptual belief) determines (in part) a GEM concept formation “path”. I propose three concept formation paths. The paths are determined by the perceptual and linguistic cues that lead to the formation of a new concept. Each path corresponds to a global category. I will refer to the three global categories as follows: (1) animate (living mover), (2) non-living mover, and (3) non-mover. I will now describe each of these in turn.

In the developmental literature, being animate is often characterized in terms of three properties: being a self-mover, exhibiting biological motion patterns, and responding contingently to events at a distance.⁷² Paradigmatic animate concepts are animal concepts such as CAT, BIRD, or HORSE. A “self-mover” is a body that can initiate and terminate its own movement.

⁷¹ I use asterisks and italics as a convention to mention the representational content of perceptual states. The characterization of the representational contents of (PS_k) and (PB_k) is a gross simplification of the representational content of any perceptual state and its corresponding basic perceptual belief. The notion of a basic perceptual belief comes from Burge 2010, pp. 546-7. (See also an earlier discussion of conceptualization of perceptual attributives in Burge 2003, pp. 545-8).

⁷² Though I am employing Jean Mandler’s characterization of an animate, I am not arguing for her empiricist account of the formation of the concept AGENT. For Mandler’s account see Mandler, 2004, pp. 94-111. For Susan Carey’s criticism of Mandler’s account of the formation of AGENT see Carey 2009, pp. 194-6.

By contrast, a “non-mover” is a body that cannot move or stop moving independently of some external force. Non-mover concepts include both non-motile artifacts such as BOTTLE and CORKSCREW, and non-motile non-artifact concepts such as TREE, FLOWER, and ROCK.

A “non-living mover” is a body that is a non-living artifact frequently seen in motion, and which does not move in the rhythmic pattern that characterizes biological motion. Vehicles such as cars, trains, and planes are non-living movers. Paradigmatic non-living mover concepts are non-living motile artifacts such as CAR, TRUCK, and PLANE.

Two important points about the three global categories are in order: First, the three global categories and their subcategories are not ontogenetically equal. Concepts that indicate animates tend to be the easiest for infants to acquire. Whereas concepts whose extensions are made up of non-self movers and are not commonly found in most households such as ROCK or MOUNTAIN tend to be acquired later. Second, my proposal is neutral with respect to the question of whether forming any concepts requires having the concepts that indicate the three global categories such as LIVING, MOVER, or ANIMATE. Thus I intend for the description of the model to be compatible with a view on which such concepts must be either innate or acquired by some other means, as well as other views on which such concepts need not be present at all. For example on some such other views the different “global paths” are built into corresponding innate perceptual input analyzers that respond to innately determined purely perceptual patterns.

Infants use both dynamic and static features to differentiate between animate and inanimate bodies. Static features include facial features such as eyes, hands, and textural features such as fur. The ability to use information from both static and dynamic features makes it possible to extract more complex information from surrounding events. Below I briefly discuss some of the experimental work that demonstrates infant sensitivity to the properties of animacy

described above. I emphasize the importance of motion cues in particular because motion perception is available very early in development. The early development of motion perception acuity means that motion is a salient feature for infants, and thus a good candidate for an early available perceptual cue.^{73, 74}

Exhibiting a biological motion pattern

“Biological motion” refers to characteristic motion patterns of animals including human beings. The motion patterns of walking in humans as well as most domestic animals can be described as “several pendulum-like motions of the extremities relative to a joint.” (Johansson, 1973) The highly regular nature of biological motion is explained by the fact that the regular pattern of movement is determined by the structure of the skeleton. As a consequence of the structure of the human skeleton, the following are characteristic properties of human motion: The limbs of the body are symmetrically aligned along a vertical axis. The limbs move in dynamic symmetry. Individual joints are rigidly connected. And joints located on the same limb share a common temporal phase of oscillation. These properties place physical constraints on the possible movement patterns of the body. Thus while biological motion is significantly more complex than mechanical motion, the regularity of the motion patterns can make it possible to

⁷³ For a comprehensive overview of infant perception see Kellman and Arterberry 2006. For examples of experiments that show infant flexibility in combining dynamic and static cues to animacy see Arterberry and Bornstein 2002.

⁷⁴ The GEM concept formation model could in principle accommodate more concept formation paths. For example, in addition to natural kinds and artifact kinds, developmental and cognitive psychologists working in the tradition of psychological essentialism have posited a form of psychological essentialism for social role kinds such as gender or race (Gelman 2003, Rhodes and Gelman 2009, Rhodes et al. 2012). I focus on the three global categories described in the text because empirical research suggests that concepts that fall under these global categories are acquired earlier in development.

use psychophysical manipulations to develop models of the heuristics used by the perceptual system in its attribution of biological motion.⁷⁵ (Berenthal, 1993, p. 178)

There are three factors that make biological motion a good candidate for an early available perceptual cue in concept formation: First, some salient biological motion is highly regular. Second, biological motion is highly prevalent in the environment both ontogenetically and phylogenetically. Third, empirical evidence suggests that infants as young as 3 months old exhibit perceptual sensitivity to biological motion, (Berenthal, 1993, pp. 182-4) and can use both static and dynamic features to distinguish between animals and vehicles (Arterberry and Bornstein, 2001). So biological motion cues could be sufficient to select the animate concept formation path early in cognitive development.

Self-Initiated Motion

Self-initiated motion is also a salient perceptual feature for infants. In a seminal experiment Alan Leslie (1982) showed that infants as young as 4 months old responded differently to animated films depicting the self-initiated movement of a block and the movement of a block upon contact with another block. In a more recent set of experiments Pauen and Träuble (2009) showed that 7-month-olds can combine information about static and dynamic features to form expectations about whether one of two objects moving together in an ambiguous motion event is a “self-mover”.

⁷⁵ A seminal research paradigm developed by Gunnar Johansson made it possible to study perceptual information based on motion patterns without the interference of information based on perception of form using a stimulus known as a “point-light walker” (1973). A Point-light walker (PLW) is made up of 10-12 point-lights organized by the location of joints on the trunk of the body and the joints of the arms and legs. A film or video recording of a point-light walker in motion (i.e. walking) is made under lighting conditions in which nothing is visible except the point-lights at the joints. When a static image of a PLW is shown, the stimulus is perceived as a randomly organized group of point-lights against a dark background. When a PLW is shown in motion subjects invariably describe the moving image as a person walking. Spontaneous and correct identification of other activities tested using PLWs include: running (in different direction and on a circular track), cycling, climbing, and dancing in couples (Johansson, 1973, p. 204, Berenthal, 1993). In more recent experiments, fMRI, neuroimaging, and single-unit studies have produced evidence that there are specialized neural mechanisms for analyzing the kinematics of biological motion (Grossman et al., 2000).

Responding contingently at a distance

The claim that infants are sensitive to contingent events between two bodies not in contact is largely based on experiments that make use of some kind of “chase” stimuli between animated shapes that generate a causal interaction illusion. (Leslie and Keeble, 1987) In a chase scenario one shape, “the chaser”, begins moving towards a second stationary shape, “the chasee”. Once the chaser is sufficiently close, the chasee begins to move away from the chaser. Experimental evidence has shown that 3-month-old infants prefer an animated chase stimulus over a random pattern stimulus. (Rochat et al., 1997). Chase scene experiments have been used to show that infants are sensitive “causation at a distance” where one the behavior of one body is contingent on the behavior of another body, with no perceptible contact between the two. (Schlottman and Surian, 1999)⁷⁶

3.2.1.2 Stage 2b: Selecting a perceptual feature space

The cues that lead to the selection of a global path also lead to the selection of what I will call a “perceptual feature space” or simply a “feature space” for short. A feature space is a cognitive structure in which perceived instances of perceptually salient features are encoded and represented.

As entries into the feature space accumulate, clusters of highly correlated features begin to emerge. Feature clusters are representations that reflect patterns in the world. The emergence of feature clusters is illustrated in figures 3.2 and 3.3 below. Figure 3.2 depicts fifteen particulars with two types of features: shape and pattern. There are three discreet kinds of shapes: square, diamond, and circle. There are also three discreet kinds of patterns: solid white, horizontally

⁷⁶ For a review of some empirical work on perception of animacy and causality see Scholl and Tremoulet, 2000. For a detailed discussion of psychophysics of generating the chase illusion see Gao, Newman, and Scholl, 2009. For a general discussion of infant sensitivity to contingency of motion between objects see Mandler 2004, pp.96-98. For seminal early experimental work on infant sensitivity to contingent events see Watson, 1972.

striped, and diagonally striped. The distribution and co-instantiation of feature kinds are encoded in the two-dimensional feature space represented in figure 3.3. We can see a feature cluster beginning to emerge in the top right corner of figure 3.3. The emerging feature cluster reflects the high correlation between being diagonally striped and being a circle. In order to depict the clustering effect I have imposed some within shape-kind and within pattern-kind variation. Without the imposed variation all of the diagonally striped circles would “stack” making it difficult to discern a clustering effect. In fact all of the diagonally striped circles are qualitatively identical, *mutatis mutandis* for the other shape/pattern combinations.

In addition to encoding statistical frequency feature spaces are a source of information for similarity judgments. Similarity can be assessed over multiple dimensions at once. The feature space depicted in figure 3.3 is a two-dimensional model that reflects similarity between particulars in terms of distance between points in a dimensionally organized metric space.⁷⁷ For the purpose of illustration I’ve supposed that the horizontal stripe pattern and the diagonal stripe pattern share a higher degree of similarity than the solid white pattern and the horizontal striped pattern. (Likewise I’ve supposed that being square-shaped and diamond-shaped share a higher degree of similarity than being square-shaped and circle-shaped). I have also supposed that similarity along both features is transitive.

⁷⁷ There are several models of similarity proposed in the empirical literature. These proposals include “geometric models” and “feature based models”. Geometric models such as non-metric multi-dimensional scaling (MDS) models represent similarity in terms of a set of points in a dimensionally organized metric space (Goldstone and Son, 2005). To take a simple example, in the case of two dimensions the similarity measure between two points will just be the shortest distance between the two points. Feature based models include set-theoretical approaches to modeling similarity such as the one proposed by Amos Tversky (1977). For a critique of geometric models see Gati and Tversky (1982). Feature based models were developed in response to evidence that similarity judgments often violate the symmetry assumption of the geometric model (Goldstone and Son, 2005). For simplicity I have employed the language of geometric models in the text.

Figure 3.2: Fifteen Particulars

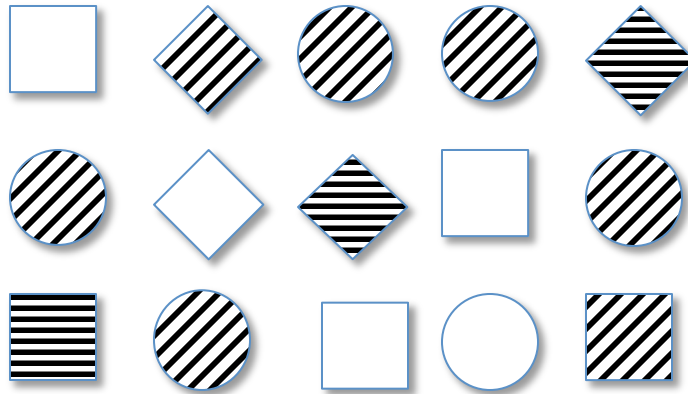


Figure 3.2: A depiction of fifteen particulars with variation along two features: shape (circle, square, and diamond) and pattern (white, horizontally striped, and diagonally striped).

Figure 3.3: A Two-Dimensional Feature Space

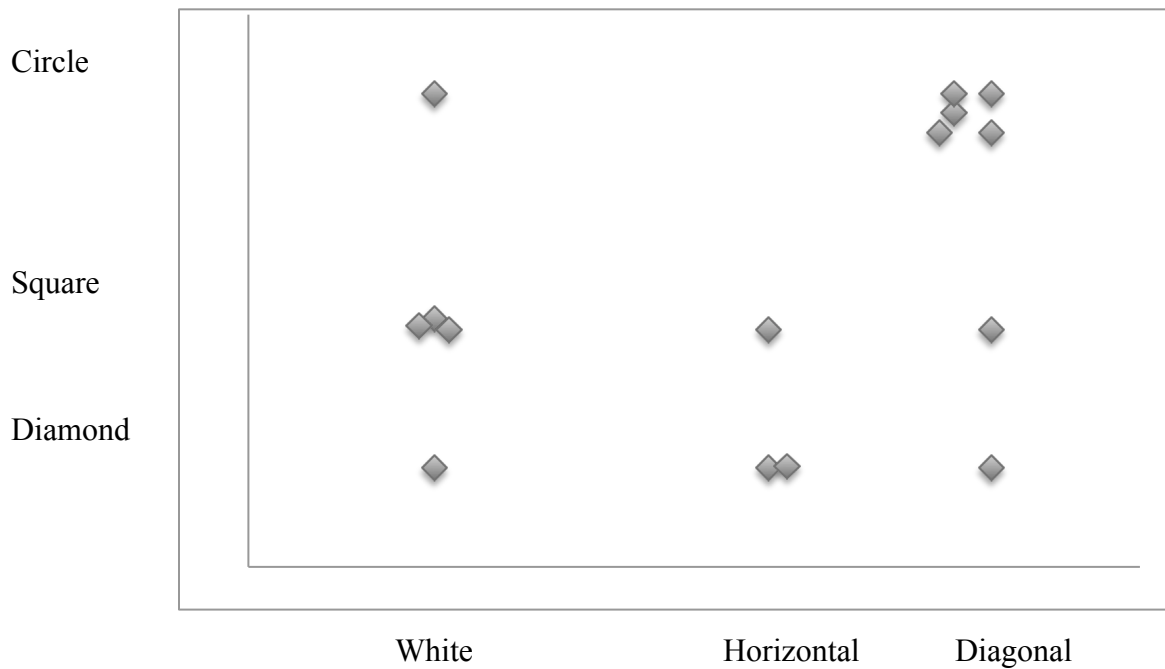


Figure 3.3: Example of a two-dimensional feature space with encoded similarity judgments. Pattern is represented on the x-axis and shape represented on the y-axis. The distance between the values on the axes represents similarity. The shorter distances represent higher degrees of similarity. In order to depict the clustering effect, some degree of within shape-kind and within pattern-kind variation among particulars has been imposed.

The particulars in the example above are qualitatively identical within shape-kind and pattern-kind. (E.g., all of the squares are qualitatively identical with respect to shape; all of the diagonally striped particulars are qualitatively identical with respect to pattern; etc.) In actuality we are perceivers in a much richer and variegated world, and our emergent feature clusters reflect highly complex patterns of correlations in the distal environment. Shape and pattern—or more realistically texture—are “low-level” features. These contrast with what are known as “high-level” features such as animal parts like eyes, tails, and wings.

Infants as young as 10 months old have been shown to form categories on the basis of both low-level and high-level feature correlations. (Younger and Cohen 1983; Younger 1985; 1990) For example, in a series of experiments, Barbara Younger tested 10-month-olds ability to form categories using stimuli that exhibited high-level feature correlation with within feature-kind variance (Younger, 1990). The correlated high-level features were head appendages (ears or antlers) and tail kind (feathered or furry). No two tails or head appendages were identical. For example, feathered tails varied in length, number of feathers, and feather shape. Likewise some antlers were wide and multi-pointed as in the case of mature male deer, others approximated two tall horns as in the case of mature male impalas. Sensitivity to correlational structure among such variant high-level features is significantly more complex than that depicted in the shape/pattern example.

3.2.1.3 Stage 3: Forming A GEM Concept

The next stage in the model is the formation of a concept. At this point in the process, the system posits a new objective kind in the distal environment. Forming a category purely in terms of feature correlation is not the same as forming a category that posits a new kind. For example, the category of things that are cats is different from the category of things that show a high correlation of typical cat features. The extension of the former includes all cats. The extension of

the latter includes cat-fakes (non-cats that exhibit the highly correlated typical cat features) and excludes deviant cats (cats that due to any number of misfortunes lack one or more of the highly correlated cat features). Forming a GEM concept on the basis of emergent feature clusters requires positing a kind that goes beyond the information present in that feature cluster.

In some cases the emergence of a feature cluster might be sufficient for the system to posit a kind and thus form a new concept.⁷⁸ The emergence of feature clusters is just one impetus for forming a new concept. Nothing in the model presented here suggests that the emergence of a feature cluster is always sufficient for forming a concept. Importantly, nothing in the model precludes the possibility that a concept is formed on the basis of perceiving a single particular either. Strictly speaking the model is only committed to the claim that at least some perceptual features are represented and that those features determine at least in part the global type of the GEM concept formed.

In the case of infants, that a concept is formed on the basis of perceiving a single particular seems most plausible for concepts with innate determiners. Alternatively, an older child or adult with extensive experience of various kinds within a global category could acquire a GEM concept on the basis of a single particular independently of innate determiners. For example, suppose you are an individual who was born and raised in North America and you have reached adulthood having never heard of capybaras. A capybara is a large South American rodent that looks (roughly) like a giant guinea pig. Just one encounter with such a beast would likely suffice for you to form the concept CAPYBARA.⁷⁹

⁷⁸ At a minimum, what it means to say that a subject S (or some part of her psychology) has formed some new concept C is that S (or the relevant part of her psychology) can now have (or include) representational states whose contents includes the representational content C.

⁷⁹ Of course a normally developed adult could also form the concept CAPYBARA simply on the basis of reading the above.

Linguistic cues are another plausible impetus for positing a new kind. For example, experimental evidence suggests that consistent linguistic labeling plays a role in category formation. (Waxman and Braun, 2005). In addition, using a new or contrasting label in presenting an object to an infant can function as a linguistic cue to posit a new kind. The results of contrastive label experiments show that using contrasting labels can help 9-month-old infants to represent objects under a kind concept. (Xu, 2002) Importantly, using contrasting tones or contrasting emotional expressions did not facilitate object individuation under a kind concept. These results suggest that the relevant cues are linguistic rather than just auditory.⁸⁰

3.2.1.4 Stage 4: Creating an associated stereotype

Once a concept is formed, a stereotype can be created and associated with that concept. A stereotype is a complex representation of the superficial features and tendencies of kinds. For example, the cat stereotype will include representations of the stereotypical motions of cats, the stereotypical colorings, textures, and shape of cats, as well as the stereotypical sounds and behaviors exhibited by cats. To be a stereotype a representation must carry some indication that the represented features in the stereotype are stereotypical *of some kind*. The stereotype associated with a concept can draw on information encoded in a represented feature cluster. The indication that the represented features in the stereotype are stereotypical *of some kind* does not have to be explicitly represented; it may be encoded in the computational role of the stereotype. Stereotypes are not constitutive of the concepts with which they are associated. They function as heuristics for defeasible concept application.

⁸⁰ Prior to these contrasting label experiments studies using the same experimental design but without the labels showed that infants' ability to represent objects under a kind concept develops between 10 and 12 months. (For further evidence of the role of naming in object individuation see Xu and Carey, 1996, as well as discussion in Carey, 2009, pp. 270-1). Kind concepts such as DOG should be distinguished from global kind concepts such as ANIMAL. Jean Mandler and her colleagues' extensive experimental work suggest that the global kind concept ANIMAL is available by 9 months and possibly as early as 7 months (Mandler 2004, Mandler and McDonough, 1993, 1998).

Stereotypes and feature clusters are importantly different kinds of representations.⁸¹ In the present model, the emergence of feature clusters is upstream of concept formation, whereas the creation of a stereotype is downstream of concept formation. As I said, a stereotype is a stereotype *of* some kind—the kind picked out by the concept with which the stereotype is associated. Feature clusters are not representations *of* any particular kind. Feature clusters may be formed from perceptions of particulars of various kinds; they are representations of frequently co-instantiated features in the world. Stereotypes and feature clusters might both project beyond the experience of the psychology in which they inhere. But the projections are different, and lead to different expectations. Stereotypes purport to cover unseen instances of the kind for which they are the stereotype of. By contrast, feature clusters only project to yet unseen *co-instantiated features* (if they project at all).⁸²

Strictly speaking, having a stereotype is not a necessary condition on having a concept. That is, one can have some concept C without creating a stereotype for the extension of C. Nevertheless, I have included stereotype creation as part of the model because creating a stereotype is part of the process of forming rough and ready heuristics for concept *application*. So while a stereotype is not part of the content of the concept with which it is associated, it can—and often does—play a role in the explanation of concept application. Note that there is no condition that says that a stereotype must be based on a large sample. Just as in the case of the capybara described above, a stereotype could be based on a single instance of some property. A

⁸¹ In Chapter 2 (and in Weber 2016) I did not distinguish between feature clusters and stereotypes, and opted for using the term ‘stereotype’ to cover both.

⁸² Experimental evidence suggests that infants do form expectations about greater populations based on samples. Fei Xu and Vashti Garcia (2008) conducted a series of experiments to test whether infants were “intuitive statisticians.” Their experiments showed that 8-month-olds form expectations about populations based on experienced samples, as well as expectations about samples based on the larger population from which the sample is drawn. For the original experiments see Xu and Garcia 2008. For a detailed discussion of the implications for Bayesian models of infant learning see Xu 2007.

stereotype is just one kind of application heuristic. With time and experience additional heuristics for concept application are created and developed.

3.2.2 The model in full

Figure 3.4 represents the concept formation model as a “decision tree” beginning with the formation of a perceptual state and perceptual belief in box 1. The tree then branches into three concept formation paths. Each path corresponds to one of the global categories described above and its respective perceptual feature space. The perceptual feature spaces that correspond to the global categories animate, non-living mover, and non-mover are represented by boxes 2a, 2b, and 2c, respectively. The first paths out of each of these boxes lead to the formation of a new concept. The second paths out of boxes 2a, 2b, and 2c (respectively) branch to representations of particulars under concepts already present in the conceptual repertoire.

From Perception to Concept Formation and Associated Stereotype Creation

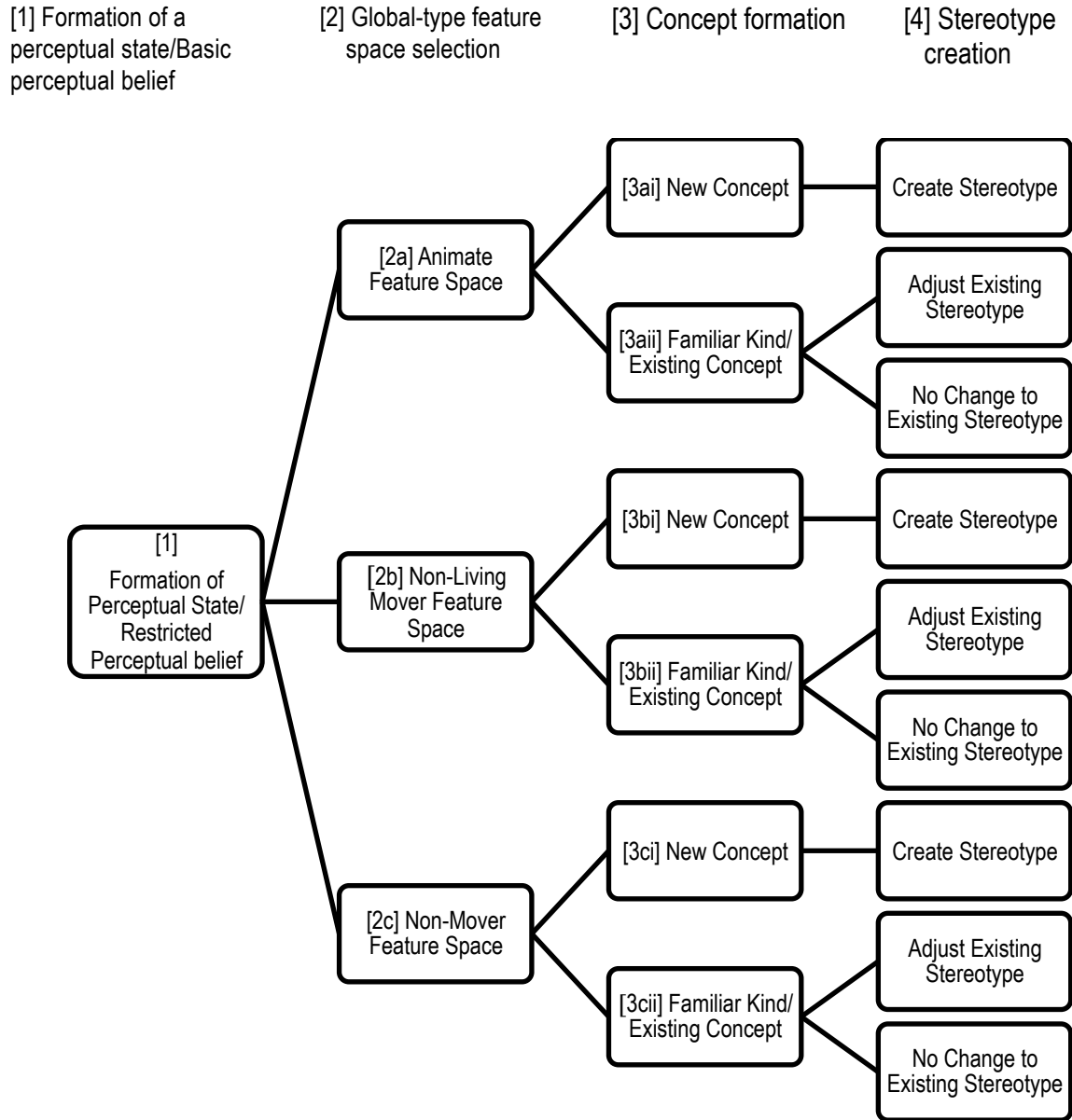


Figure 3.4: GEM concept formation followed by stereotype creation depicted as a tree summarizing the different possibilities at each stage described above. For simplicity the above diagram leaves out cases in which a new concept is formed but no stereotype is created.

3.3 Part III: The Place of Concept Formation in the Epistemology of Empirical Discovery

Now that we have the basics of a model of GEM concept formation laid out, let us consider the place of concept formation in the epistemology of empirical discovery in more detail. Recall that the goal of empirical discovery is warranted empirical belief, and ultimately, to achieve an understanding and knowledge of the empirical world. Concept formation is how one gets the epistemic endeavor of empirical discovery off the ground. Forming concepts is the first step in making commitments about the empirical world, where the contents of those commitments go beyond the representational contents of perceptual states and basic perceptual beliefs. Epistemic theories rarely discuss the norms for transitioning from basic perceptual beliefs to general empirical beliefs where that transition requires forming a new concept. Such transitions require using (and thus having) general concepts; concepts whose content goes beyond that of perceptual attributives that indicate such properties as color, shape, and body. We need such general concepts to engage in empirical investigation.

The GEM concept formation model presented in part II shows how concepts can be formed using very little information. The biases described in the model make the problem of forming concepts tractable. However, that same poverty of information makes it difficult to see how using concepts in our earliest beliefs can be warranted. The problem is brought into focus when we consider the infant's impoverished representational repertoire. How can she be warranted in forming beliefs about that which she knows almost nothing? I claimed that our warrant for forming beliefs about an empirical subject matter of which we know almost nothing, is inextricably linked to our ability to form new ways of representing the world. Infants provide the most intuitive example, but adults are subject to the same epistemic circumstances. We encounter items with which we are entirely unfamiliar, and in the face of ignorance we begin to

form beliefs. (Recall the example of the capybaras described in part II). On my view we are warranted in doing so at least in part by our entitlement to apply newly formed concepts.

On my proposal, the prima facie entitlement (and thus prima facie warrant) achieved in forming GEM concepts on the basis of perceptual experiences has two sources: (1) the nature and function of our sense-perceptual capacities by which perceptual states and perceptually derived feature clusters are formed; and (2), the contribution that the process makes to the fulfillment of epistemic norms that govern the formation of empirical beliefs and empirical discovery. In Chapter 2 I suggested one such set of norms that I call “revisability norms”. A belief is empirically revisable if there is an epistemically warranted route to the rejection of the belief, given some possible empirical evidence. Revisability contributes to securing long-run reliability of our fallible belief forming processes. Revising some belief (Bp) requires new information that bears on the question of whether or not p .⁸³

In what follows I re-describe the GEM concept formation model as a normative model that captures epistemic standards for forming early beliefs using newly formed concepts. The normative model is made up of four defeasible presumptions that correspond to the four parts of the GEM concept formation model as described in Part II. The defeasible presumptions make it possible to start forming warranted beliefs in the face of tremendous ignorance. They support the transition from basic perceptual beliefs to beliefs that require the formation and application of new concepts. The normative model illustrates the relationship between GEM concept formation, early empirical belief formation, and warranted empirical belief revision. All three jointly contribute to empirical discovery. I will review each part of the model, its corresponding defeasible presumption, and the types of errors that are a consequence of presumption failure.

⁸³ Revisability is contrasted with mere replaceability. A belief is replaceable if it can be overwritten by some other new belief. Replaceability can be brute (i.e., achieved by taking a pill, getting hit on the head, etc.)

3.3.1 The defeasible presumptions that back GEM concept formation and early application

3.3.1.1 The Veridical Perception Presumption

Recall that GEM concept formation begins with the formation of a perceptual state and a basic perceptual belief on the basis of that perceptual state. The basic perceptual belief is a singular belief that purports to pick out a particular body at a location and predicates attributes that correspond to perceptual attributives such as ***yellowish***, ***furry***, ***round***, as well as attributives that specify motion patterns. The stage 1 defeasible presumption can be formulated roughly as follows:

Presumption 1: “Veridical Perception Presumption”

The perceptual system is reliable in the present environment and it produces perceptual states that are veridical.

Reliability in Presumption 1 is relative to the normal environment, that is, the environment in which the content of the perceptual system is explained and established. Thus presumption 1 includes the general presumption that the individual is in the normal environment.⁸⁴

When conditions are such that the veridical perception presumption fails, false perceptual beliefs can occur. For example, this is what happens when false perceptual beliefs are based on visual illusions. Such illusions can be referential illusions or attribution illusions. In the case of a visual referential illusion, a non-veridical perceptual representation as of a body is formed on the basis of proximal stimulation when in fact there is no body causing the proximal stimulation. For example, consider the case of the perceptual experience as of a figure approaching in the fog, when in fact there is no body approaching. In such cases one has the illusion that there is a body

⁸⁴ See argument in Burge, 2003, pp. 532-6 for detail, and a review of that argument in Chapter 2 of this dissertation. I return to the general presumption of normality after reviewing the other defeasible presumptions.

approaching due to the refraction of the light through water crystals in the fog.⁸⁵ In the fog case the corresponding revisable false empirical belief is:

(B1) THAT IS A BODY THERE.

In the above example (B1) is a basic perceptual belief based on the perceptual state (PS1) with the representational content ***that body there***⁸⁶. The instance of THAT in (B1) fails to refer.

(B1) is an instance of a revisable empirical belief that is a consequence of presumption 1 failure.

The perceptual system may very well be reliable in the present environment, but the perceptual state (PS1) is not a veridical state. (B1) is revisable because there is an epistemically warranted route to the rejection of the belief, given some possible empirical evidence. The empirical evidence could be a perceptual belief based on a veridical perceptual state formed immediately after (PS1) and then reaffirmed by more careful investigation.

In the case of a visual attribution illusion there is a particular body that is the distal source of the proximal stimulation that causes the formation of a perceptual state, but something in the ambient conditions causes an attribution error. For example, a white globe being lit by a red light will appear red under certain conditions. When such conditions obtain, the result will be a non-veridical perceptual state that leads to the following false basic perceptual belief:

(B2) THAT IS A RED BODY THERE.

In the case of (B2) the demonstrative THAT succeeds in picking out a body at a location, but (B2) is false because the body is in fact white and not red. As in the case of (B1) the perceptual system produced a non-veridical perceptual state which, when conceptualized, yielded a false empirical belief. (B2) is revisable in the same way that (B1) is revisable.

⁸⁵ For an extended discussion of referential illusions see Burge 2010a, pp. 382-90. The example is inspired by the discussion of referential illusion in Burge, 2009, pp. 291-3.

⁸⁶ I have simplified the representational content of (PS1) and (B1) for the sake of exposition.

It should be noted that the veridical perception presumption is a component of our perceptual entitlement to hold our basic perceptual beliefs. Perceptual entitlement for basic perceptual beliefs is part of the explanation of concept formation entitlement because the non-accidental reliability of the perceptual system is part of what makes the GEM concept formation process a source of warrant.

3.3.1.2 The No Global Fakes Presumption

Recall that the next stage in GEM concept formation is determining a concept formation path and thereby selecting an associated feature space for recording salient perceptual features. Recall that the concept formation path is a consequence of perceptual and linguistic cues which bias the concept formation system into selecting one perceptual feature space over the others. The paths and feature spaces correspond to defeasible global classifications of the downstream concept being formed. The corresponding defeasible presumption can be formulated roughly as follows:

Presumption 2: “No Global Fakes Presumption”
There are no cross-global-type fakes.

The no global fakes presumption underwrites the defeasible classification into one of the three global categories animate, non-living mover, and non-mover. Cross-global-type fakes exhibit the superficial features that result in perceptual cues that cause the selection of the wrong concept formation path and associated feature space.

An example of a cross-global-type fake is the Putnamian robocat described in Chapter 2. In Putnam’s robocat scenario all of the entities on Earth that we think of as cats turn out to be robots controlled by radio from Mars (Putnam, 1962). A typical robocat looks and acts just like a cat. Thus a perception of a robocat will produce the same perceptual cues as a veridical perception of a cat, and will thereby cause the selection of an animate (living mover) concept

formation path, thereby miscategorizing the robocat at the global level. The robocat is wrongly categorized at the global level because robocats are actually non-living movers.

The robocat scenario is a scenario in which presumption 2 fails. Examples of false revisable empirical beliefs that are a consequence of presumption 2 failure in a robocat scenario include:

(B3) THOSE ARE BIOLOGICAL ANIMATES (pointing at robocats).

(B4) ROBOCATS ARE ANIMALS.⁸⁷

(B3) and (B4) are empirically revisable false beliefs, so there is a possible epistemically warranted route to the rejection of the beliefs, given some possible empirical evidence. For example, in this scenario scientists could develop the technology to do microscopic surgery. They might then discover that upon closer inspection, what look like blood vessels in robocats are, in fact, cleverly disguised wires that connect to tiny circuit boards. The evidence that these creatures are not made of biological materials would provide an epistemically warranted route to the rejection of (B3) and (B4).

The robocat scenario is a case in which a non-living mover is mistaken for an animate. We can easily construct other cases, cases in which an animate is mistaken for a non-mover, or a non-living mover is mistaken for a non-mover. Arguably the movie *Toy Story* is full of examples of the latter. In the movie, what are seemingly non-moving toys such as a piggy bank and a Mr. and Mrs. Potato head, turn out to be movers. Though the movie doesn't specify whether the toys are in fact non-living movers or animates, we can easily fill out the example for either case.

⁸⁷ I use the concept ROBOCATS because that is the concept that is formed given that there are no cats in the scenario described above. If the word 'cat' is used to express the concept ROBOCAT then the following false belief could also arise: (B4*) THE CREATURES WE CALL 'CATS' ARE ANIMALS.

3.3.1.3 The One-Look-One-Kind Presumption

Once the perceptual information has been recorded in one of the three global feature spaces, the concept formation system must “decide” whether the particular generating the present perceptual state is an instance of a new kind or an instance of a previously experienced kind. The presumption that supports the formation of a new concept at this stage is the following:

Presumption 3: “One-Look-One-Kind Presumption”

There is one kind for every feature cluster. An unfamiliar looking particular is an instance of a new kind, and a familiar looking particular is a particular of a previously observed kind.

An example of a one-look-one-kind presumption failure occurs in what I’ll call “the cragon world” scenario: Cragons are small (approximately cat-sized) dragons that regularly disguise themselves by wearing realistic-looking cat costumes and adopting the mannerisms and behaviors of cats. In the cragon world both cragons and cats exist; they are distinct natural kinds that have the same look. Suppose you’ve seen many cats but just happen to come across your first cragon. This is a case where a familiar looking particular is in fact an instance of a previously unobserved kind.⁸⁸ The one-look-one-kind presumption pushes you towards believing that there is only one kind with that familiar look. So your first sighting of a cragon is decidedly unremarkable because you assume that you are seeing yet another cat. Both the perceptual system and the concept formation system process the perceptual experience of a cragon as a perceptual experience of the familiar kind. No new concept will be formed. The one-look-one-kind presumption failure in the cragon world could lead to the following false beliefs:

(B4) THAT’S A CAT (pointing at a cragon).

(B5) THAT ANIMAL CAN’T BREATHE FIRE (pointing at a cragon).

⁸⁸ A real world example of a one-look-one-kind failure is the case of jadeite and nephrite (recounted in Putnam 1975). Jadeite and nephrite are two different minerals that look very similar and were thought to be one kind of material known as jade until the differences in their chemical and physical properties were discovered. (Putnam, 1975/1975 p. 241)

A particular is identified as having a “new look” or “an old look” by consulting the present state of the feature clusters represented in the perceptual feature space. In this case there is no “new look”, but that is because cragons are cat fakes. As in the case of robocats, given some empirical evidence, there is an epistemically warranted route to the rejection of (B4) and (B5). For example, due to some unfortunate accident, one day a cragon’s custom might start to give or tear and thereby reveal the cragon’s true scaly exterior.

3.3.1.4 The Fakes-And-Deviants-Are-Rare Presumption

Recall that once a new concept is formed an associated stereotype can be formed. The stereotype functions as a heuristic for concept application and defeasible categorization of particulars. The presumption that underwrites the formation of the stereotype is the following:

Presumption 4: “Fakes-And-Deviants-Are-Rare Presumption”

The sample on the basis of which a stereotype is formed is not made up of fakes or deviants.

A stereotype is supposed to be a heuristic for identifying and drawing inferences about unseen instances based on the sample that has been previously observed. Thus there is a presumption that the sample on the basis of which the stereotype is formed is in fact representative of unseen instances.

As formulated, the no global fakes presumption and the fakes-and-deviants-are-rare presumption include some redundancies. This is to be expected because fakes and deviants can lead to presumption failures both across global categories and within global categories. A global-type fake is a fake by definition so overlap with the no global-type fakes presumption is to be expected.

The fakes-and-deviants-are-rare presumption follows from the one-look-one-kind presumption. If there is one kind for every feature cluster and an unfamiliar looking particular is an instance of a new kind, then an unfamiliar looking particular is not a deviant instance of a

previously observed kind. And if a familiar looking particular is a particular of a previously observed kind, then it is not an instance of a new kind that has the look of a previously observed kind, (i.e., it is not a fake). If this is true of all the particulars that have been encountered, then the stereotype is formed on the basis of a sample that is not made up of fakes or deviants. That is, if presumption 3 holds, then presumption 4 holds as well.

I call errors that are a consequence of the *fakes and deviants are rare* presumption failure “typical-atypical errors”. Whenever there is a one-look-one-kind error there will also be a typical-atypical error. This is the case because one-look-one-kind presumption failures and fakes-and-deviants-are-rare presumption failures go together. Still, it is important to distinguish these error types because they reflect an important difference between the formation of a concept and the formation of a stereotype. The fakes-and-deviants-are-rare presumption underwrites the utility of forming stereotypes at all. If the world were riddled with fakes, then stereotypes would not function as useful heuristics for defeasible categorization. Moreover, if deviants were not rare, then there would be no stereotypical characteristics in the first place. Deviants must be rare because if there were no stereotypical characteristics, then there would be no stereotype from which to deviate.

Typical-atypical errors occur in cases where the particulars that make up our sample of some kind turn out to be atypical members of that kind. Consider the following scenario I call “the valley of orange oranges” (inspired by Putnam, 1970/1975): Suppose that up until this moment, all of the ripe oranges you have ever seen have been orange colored and spherical shaped. Unbeknownst to you, you live in an anomalous valley of orange oranges. As it turns out, most of the world’s ripe oranges are actually green and oblong. If you were to see what is in fact a typical orange, you would not recognize it as such. The combination of the one-look one-kind

presumption and the fakes-and-deviants-are-rare presumption would lead you to believe that you are seeing a particular of a previously unobserved kind. In this scenario you form the following false empirically revisable beliefs:

(B6) THAT IS NOT AN ORANGE (when pointing at a green orange).

(B7) MOST ORANGES ARE ORANGE.

As in the previous cases, there is an epistemically warranted route to the rejection of (B6) and (B7) given some possible empirical evidence. For example, DNA analysis of deviant orange oranges and typical green oranges could reveal that they are the same species. Furthermore, it could be discovered that there is a type of gas present in the valley of orange oranges that prevents oranges from developing in the way they would in any environment outside of the valley. This additional evidence would explain why these deviant oranges are typical within the confines of the valley.

3.3.2 The role of the defeasible presumptions in concept formation entitlement

Now that we have a detailed account of the defeasible presumptions that capture the GEM concept formation model, I will explain the relationship between the defeasible presumptions and concept formation entitlement.

The defeasible presumptions make it possible to start forming warranted beliefs in the face of tremendous ignorance. In Chapter 2 I argued that being warranted in acquiring a concept *C* by some process *M* yields a prima facie entitlement to apply *C* in forming empirical beliefs. To yield that prima facie entitlement *M* must produce concepts in a way that is non-accidental relative to our representational capacities. Concept formation entitlement is a kind of warrant that can attach to concepts formed on the basis of perceptual experience including perceptually derived representations such as perceptual feature clusters.

As I said in Chapter 2, concept formation entitlement is a kind of entitlement that is a species of the Burgean notion of epistemic warrant. Epistemic warrant is an epistemic good that is achieved by a psychology's meeting norms or standards that contribute to the fulfillment of the representational function of forming or preserving true beliefs. Entitlement is warrant without reason. It is warrant that derives purely by a psychology's operating well and reliably with respect to its representational functions (Burge, 2013). So concept formation entitlement does not require that the defeasible presumptions be represented in an individual's psychology.

The defeasible presumptions support the transition from basic perceptual beliefs to beliefs that require the formation and application of new concepts. This is the transition that concept formation entitlement is meant to explain. The presumptions support that transition in at least two ways: First, they reliably track real patterns in the world. And second, they guarantee the revisability of our empirical beliefs. I'll now consider each of these ways in turn.

3.3.2.1 The defeasible presumptions reliably track truth

While the warrant that derives from our perceptual and conceptual formation systems depends on *metaphysically contingent* patterns in our world, the presumptions that track these patterns are overwhelmingly likely to be true. Biases based on perceptual cues that determine the selection of the concept formation path track truth and deeper structure because of the regularity of the relationship between perceptually accessible properties and GEM properties. That is why perceptual entitlement is a component of concept formation entitlement.

If the perceptual cues did not track real patterns, a reliable perceptual system that regularly produced veridical perceptual states would not be a good guide to deeper structure. For example, imagine a world in which that biological and non-biological motion patterns did not track biological and non-biological kinds. Suppose that the biological motion pattern is instantiated by only half of animals, and which animals exhibit the biological motion pattern is

not species specific. So, for example, despite the fact that all cats are biological creatures, some cats move in a way that exhibits a biological motion pattern and some cats move in a way that exhibits a mechanical motion pattern. Suppose also that in this world some non-living movers exhibit the biological motion pattern. In such a world, a perceptual system could still be reliable in attributing biological motion, but exhibiting a biological motion pattern would only track the property of being a mover that exhibits the biological motion pattern. So in this world motion patterns would not be a good guide to tracking animals. Moreover, the biological motion pattern would not be good basis from which to make inductive generalizations about deeper structure.

By contrast, in our world the biological motion pattern cue is a useful guide to deeper structure because it successfully tracks a category (animal) that is inductively richer than the category things-that-exhibit-the-biological-motion-pattern. A category is inductively rich if it can be used to generate lots of ontological commitments that contribute to the understanding of the structure of the world.

All of the presumptions are ultimately underwritten by a general presumption of normality. Each presumption is a claim about the normal environment and a presumption that the present environment *is* the normal environment. In the normal environment the perceptual cues described in Part II are good guides to tracking truth and deeper structure. The normal environment is the one in which the biases shaped by the perceptual cues are explained and established. Forming a new concept in an epistemically good way requires that the process be non-accidentally reliable in the normal environment. A non-accidentally reliable process must be grounded in the representational nature of the psychological capacities that make up the concept formation process, and the constitutive relation between those capacities and the capacity for forming empirical beliefs.

Just as the biases track truth and deeper structure, the defeasible presumptions track truth and deeper structure in the normal environment. The concept application in early belief formation proceeds in accordance with the defeasible presumptions. The presumptions in turn reflect the GEM concept formation process.

3.3.2.2 The defeasible presumptions secure the possibility and set the limits of warranted belief revision

A secondary way in which the defeasible presumptions support the formation of warranted early empirical beliefs is by securing the possibility of warranted belief revision. If the presumptions that corresponded to each part of the concept formation process were not defeasible, then the concepts produced by the process would carry non-defeasible ontological commitments. For example, consider the following false belief in a robocat world: (B8) ROBOCATS ARE LIVING CREATURES. If there were no epistemic possibility of global-type fakes, or global categorization errors, the concept produced from the experience of seeing a robocat in a robocat world would carry with it the unrevisable commitment that robocats are living movers. In such a case there would be no warranted route to empirically revise (B8) given some empirical evidence.

Forming concepts that indicate kinds that can be the subject matter of empirical beliefs, and to do so in a way that ensures the possibility of a warranted route to true beliefs given some empirical evidence is how empirical discovery gets off the ground. In our investigations of the empirical world that warranted route often winds through false beliefs and belief revision.

Recall the concept formation “decision tree” diagram in figure 3.4 above. Think of the nodes of the “tree” as empirical beliefs whose subject matter is the properties indicated by GEM concepts. These empirical beliefs are revisable in so far as there is a corrective path back from a

node to a different node by abandoning some defeasible presumption. Though (B1)-(B8) are false, they could be warranted false beliefs in the scenarios described above.

The defeasible presumptions also explain the limits of empirical belief revision. At the limits of belief revision are instances of what I call “maximal revision”. Maximal revision is made possible by there being a path from any empirical belief back to a defeasible *de re* belief—the basic perceptual belief at the beginning of the GEM concept formation process. For example, suppose that the myth of the Loch Ness monster originated from a perceptual illusion as of a creature in the lake (a.k.a “the loch”) as seen through fog. Suppose that over time there were many such referential illusions experienced by different people and which led to the positing of a loch ness monster kind.⁸⁹ In this scenario an empty GEM concept is created on the basis of false basic perceptual beliefs. The concept LOCH NESS was formed on the basis of referential illusions.

On my account the early beliefs about loch ness monsters can be warranted. This result is correct even if it does not accord with our intuitions. Assuming of course that there is no evidence to the contrary, the individual (or group of individuals) in this scenario has formed a concept and beliefs on the basis of several perceptual states that purport to pick out particulars. Though they are false the beliefs are empirically revisable. We can understand what it is for it to turn out that our whole theory of what loch ness monsters are is moot. The theory is moot because it was formed on the basis of particulars that never existed. This is different from a case in which the theory was formed on the basis of veridical perceptions of bodies that turned out to not be loch ness monsters. The case in question is one in which there were no bodies out in the pond to begin with.

⁸⁹ In this example I am supposing that there is a general concept LOCH NESS MONSTER, as opposed to a singular concept such as NESSIE that purports to pick out a particular affectionately named ‘Nessie’. Recall that a general concept indicates a property or a kind as opposed to a particular. Though in fact ^loch ness monster^ has never been instantiated, we can imagine that multiple individuals believe that they have seen such creatures in the lake. And that with time the folk surrounding the lake have come to believe that the lake is home to a whole pod of such creatures.

If the veridical perception presumption were not defeasible, then we could not make sense of the claim that it turns out that there are no and never have been any particulars (lochness or otherwise) on which our theory is based. In the case of empirical beliefs, belief revision can only go back as far as beliefs about the existence of particulars.

The defeasible presumptions, and the general presumption of normality, are part of the explanation of the phenomenon of massive empirical dubitability: The ability to consider and bring into doubt nearly all of our empirical beliefs (though not all at once).⁹⁰ Throwing empirical claims into doubt merely requires conceiving of some piece of empirical evidence that can launch the unraveling of our empirical commitments.⁹¹

3.4 Part IV: Reflections and conclusions

The empirical model of Part II shows that we can form a concept that indicates some objective property despite having almost no information about that property. We do this using very little information from our immediate environment. We can do so because of the biases that make up the concept formation system. That we form concepts under conditions of such informational poverty makes it difficult to see how early beliefs that make use of such concepts could be warranted. Concept formation entitlement is warrant achieved purely by our psychology's operating well and reliably with respect to its representational function.

⁹⁰ The phenomenon of massive dubitability is highlighted by Burge in his reflections on, "the intellectual norms which provide the permanent possibility of challenge to any actual practices of individuals or communities that we could envisage." (1986/2007b, p. 274)

⁹¹ The reader may agree with the above examples of belief revision, but argue that the normative model could not plausibly apply to concepts such as ROCK or PUDDLE. Coming up with a thought experiment that applies to rocks seems plausible. To begin with, we can imagine a world in which it turned out that rocks were actually very timid animates that do not move in the presence of other animates. Coming up with a thought experiment that applies to puddles requires a considerably more complex piece of science fiction. But more important than the details of any such story is the fact that questions that challenge our empirical commitments with respect to puddles can be posed. Puddles are objectively real. The belief that there are puddles carries ontological commitments about the empirical world. We understand what it is to put our empirical commitments to the test. That ability extends to our empirical commitments about rocks, puddles, and much more.

The foregoing discussion has been an attempt at explaining the link between our ability to form warranted beliefs about empirical subject matter of which we know almost nothing, and our ability to form new ways of representing the world.

Before concluding this chapter I want to reflect on the relationship between the concept forming process, the defeasible presumptions described above, and the characterization of concept formation entitlement from Chapter 2.

In Chapter 2 I said that being warranted in acquiring a concept *C* by some process *M* yields a prima facie entitlement to employ *C* in forming empirical beliefs. That entitlement is at least part of the explanation of the warrant of beliefs that employ newly formed concepts in the face of tremendous ignorance. I argued that warrant requires forming concepts in a way that is non-accidental relative to our capacities. In Part III of the present chapter I said that the defeasible presumptions support the transition from basic perceptual beliefs to beliefs that require the formation and application of new concepts.

At this point, I can imagine an objector raising the following objection: The defeasible presumptions are doing all the heavy epistemic lifting. Why aren't the defeasible presumptions by themselves sufficient to warrant early application of newly formed concepts? If the defeasible presumptions both track truth and preserve the revisability of empirical beliefs, then the process by which we form new concepts might not matter. Rather, all that matters is having the right presumptions which track real patterns in the empirical world—the world that is the subject matter of empirical beliefs and empirical discovery.

The objection suggests that regardless of the origin of your concepts, having concepts along with the defeasible presumptions leave you well positioned to accumulate warranted empirical beliefs about the world. The objector claims that—regardless of how you acquired the

concept CAT—if you form beliefs in accordance with the defeasible presumptions and use the latest scientific instruments, you will be perfectly well equipped to carry out a successful cat research program. Depending on the situation vis-à-vis cats in your world, you could find out any number things: You could find out that cats do not typically look or behave in the way that you supposed. You could find out that some of the things you thought were cats are actually robots and some are sly dragons masquerading as cats. You could discover that some of the things you did not recognize as cats are actually cat deviants. And you could understand the errors you may have made with respect to at least most of your cat hypotheses. If the presumptions can secure the long-run reliability of your cat beliefs, how is the concept formation process contributing to the warrant of your cat beliefs?

I take it that the objector is suggesting that while we obviously need some way of forming concepts in order to have concepts, once we have concepts the presumptions are what underwrite the warrant of early beliefs in the face of ignorance. The objection grazes the surface of a deeper question. The deeper question is about the relationship between the empirical model presented in Part II and the normative account in Part III. Specifically, what is the significance of the parallel between the concept formation process and the defeasible presumptions that secure the epistemic possibility of warranted empirical belief revision?

Each presumption and the errors produced by presumption failure corresponded to a stage in the concept formation and stereotype creation process. That the presumptions capture the biases of the concept formation and stereotype creation process is neither accidental nor explanatorily inert.

Recall that I proposed that the prima facie entitlement (and thus prima facie warrant) achieved in forming GEM concepts on the basis of perceptual experiences has two sources: (1)

the nature and function of our sense-perceptual capacities by which perceptual states and perceptually derived feature clusters are formed; and (2), the contribution that the process makes to the fulfillment of epistemic norms that govern the formation of empirical beliefs and empirical discovery. The presumptions were offered as part of the explanation of how revisability norms are satisfied. But to do the relevant explanatory work the presumptions must be understood in terms of the first source of entitlement—that is—in terms of our representational capacities.

Understood independently of our representational capacities and representational success, the presumptions can only contribute to long-run reliability accidentally. When considered independently of the process for forming concepts, the presumptions are just heuristics for forming beliefs that just happen to secure the long-run reliability of beliefs in our world.

To get at the idea consider the following heuristic: “Go with how things seem.” “Go with how things seem” might be good advice for all sorts of reasons under various conditions. But it is only good *epistemic* advice if how things seem tends to non-accidentally track how things are. That the representational system is non-accidentally onto real patterns in the world is ultimately what makes the presumptions good epistemic advice.

The power of the four defeasible presumptions is secondary to the general presumption of normality. The presumption of normality is only helpful because the “normal conditions” that are presumed are the conditions in which the concept formation biases are established and explained. The biases track real patterns by exploiting perceptual cues, and therein by exploiting our sense-perceptual capacities.

Conclusion

My case for the thesis that concept formation is epistemically evaluable began with a simple intuition: Concept formation on the basis of perceptual experience often looks like a transition that is an instance of doing well with respect to some representational standards. Empirical concepts are commonly formed on the basis of experience with typical instances of the properties they indicate. That is, for example, a good way to get someone to form the concept CORKSCREW is to arrange the world such that she has lots of experiences with typical instances of corkscrews. Similarly, if you want someone to form the concept BLUEBERRY, you would be well advised to provide her with typical instances of blueberries. The fact that people often form an empirical concept on the basis of experience with typical instances of the property indicated by that concept is why concept formation looks like an instance of doing well by epistemic standards. As we saw in Chapter 1, even Jerry Fodor—who ultimately concludes that concept formation is a brute-causal neurological process—concedes that concept formation has the look of rational relation. (1981, p. 275)

Moving beyond the intuitive reasons for thinking that concept formation is epistemically evaluable, I situated concept formation within the greater epistemic endeavor I called “empirical discovery”. The aim of empirical discovery is to form warranted empirical beliefs, and ultimately, to achieve understanding and knowledge of the nature and structure of the empirical world.

I explained that understanding of the structure of the empirical world cannot be achieved by simply collecting truths about how a variety of particulars have behaved under various conditions at some time. The challenge of discovery is that to learn and understand empirical truths is not merely to compose and confirm a list of hypotheses and attendant theories.

Discovery requires creating the representational tools from which such hypotheses and theories are composed. Empirical concepts are some of these representational tools. Thus forming empirical concepts is part of the greater epistemic endeavor of empirical discovery.

The pursuit of empirical knowledge begins with the formation of perceptual states and basic perceptual beliefs. Recall that a basic perceptual belief is a conceptualization of a perceptual state. Conceptualization is a process by which perceptual representation is converted into propositional representation. Thus, a basic perceptual belief can only attribute properties that are represented in perception such as motion pattern, shape, color, texture, location, and the property of being a body.

If our representational toolbox contained only perceptual attributives, we could not form beliefs about properties other than those represented in perception. In order to form beliefs about properties not represented in perception, we need to form concepts that go beyond perceptual attributives. In Chapter 2 I gave the example of the botanist who discovers tomatoes. Let us change the example to a young child who has happened to stumble upon the first tomato ever seen in her community. Upon seeing the beautiful red ripe tomato hanging on the vine the child forms the perceptual belief (B1) THAT IS A RED ROUND BODY THERE. But the child cannot represent that red round body *as* a tomato before she forms the concept TOMATO. Thus, in order to form the belief (B2) THAT IS A TOMATO, the child must form a new concept. This is an example of a transition from a basic perceptual belief to an empirical belief that requires employing (and thus having) a general concept. The representational content of the empirical belief that represents something *as* a tomato goes beyond the perceptual attributives of perceptual beliefs. We need such post-perceptual general concepts, such as TOMATO, to engage in empirical investigation. Forming perceptual states is the first step in making empirical

commitments about the world. Forming concepts is the first step in making empirical commitments that go beyond the commitments of perceptual states and basic perceptual belief.

Intuitively, concept formation on the basis of perceptual experience has the look of an epistemic transition, and a place within an epistemic project. However, concepts cannot be true, false, veridical, or non-veridical. So concept formation is a psychological process that cannot be explained by a simple analogy to perceptual state formation or belief formation. Moreover, as I explained in Chapters 1 and 2, concept formation on the basis of perceptually derived representations is subject to what I called “the projection problem.” I described one kind of perceptually derived representation that I called a “feature cluster.” Recall that feature clusters are representations that reflect patterns of highly correlated features in the world. For example, a feature cluster that emerges from perceptual experience with cats will represent the stereotypical properties of cats as highly correlated features.

The projection problem arises because forming a concept on the basis of a representation of frequently co-instantiated features is an instance of an underdetermined transition between psychological representations. The representational content of the first representation underdetermines the representational content of the second.

To overcome underdetermination, we need an explanation of why, for example, the cat feature cluster leads to the formation of the concept CAT rather than some other concept that indicates a property whose instance could produce an identical feature cluster. The projection problem arises both for co-extensive concepts such as the concept GRANNY’S FAVORITE ANIMAL and concepts that indicate look-alikes such as CRAGONS and PUTNAMIAN ROBOCATS.⁹²

⁹² Recall that in Putnam’s robocat scenario robocats look and act just like cats but are actually all robots controlled by radio from Mars (Putnam, 1962). Cragons, on the other hand, are small (approximately cat-sized) dragons that regularly disguise themselves by wearing realistic-looking cat costumes and adopting the mannerisms and behaviors of cats.

I introduced the projection problem as part of Fodor's argument for what I called his "new radical conclusion." According to Fodor's new radical conclusion, concept formation is not a cognitive process, does not admit of psychological explanation, and cannot be explained at the intentional level. Fodor claims that the projection problem shows that the transition between feature clusters and concept formation is not rational or inferential.⁹³ He draws this conclusion because, as he says, there are "so many different kinds of relations in which a concept and [a feature cluster] may stand." (2008, p. 154) According to Fodor, this shows that there is no "pattern of inference which, given just a [feature cluster], identifies the corresponding concept." (2008, p. 154)

I did not argue for the claim that concept formation is a rational process. I take this to be a stronger claim than the claim that the process is epistemically evaluable or can be warranted. I noted that—*pace* Fodor—the underdetermination of concepts by feature clusters does not straightforwardly show that the transition is not epistemically evaluable. Nor does it show that the transition fails to meet epistemic standards. Inductive and abductive inferences are instances underdetermined transitions that can meet epistemic standards. Still, the projection pattern raised the following question: What pattern of psychological transitions can explain why a feature cluster leads to the formation of one concept rather than some other concept?

To answer this question we needed an example of a concept formation model with a regular non-ad-hoc pattern of transitions from feature clusters to the newly formed concept. I offered the GEM concept formation model as an empirically informed account of the regular pattern of psychological transitions that make up the GEM concept formation process.

⁹³ For continuity within the conclusion I am using the expressions 'concept formation' and 'feature cluster' instead of 'concept acquisition' and 'stereotype'. For a note on why my terminology differs from Fodor's see the introduction to the dissertation and footnote 21 in Chapter 2.

Recall that the acronym ‘GEM’ is derived from the three characteristics that unify the paradigmatic concepts that the model covers. GEM concepts are *general* concepts that indicate objective properties that are instantiated in whole particulars that can be picked out in perception. Such properties are the subject matter of *empirical beliefs*. The particulars in which such properties are instantiated tend to be *medium-ish* in size. Mediumish objects are perceptually discriminable independently of theory (though not necessarily independently of some instrument).

The model is a model for GEM concept formation *and* stereotype creation. It begins with the formation of a perceptual state and a basic perceptual belief. Based on perceptual cues such as motion patterns, a concept formation path is selected. I described three kinds of concept formation paths that correspond to three global categories. The categories are animates, non-living movers, and non-movers. Examples of animate concepts include animal concepts such as BIRD, CARIBOU, and KOALA. Examples of non-living mover concepts include concepts of artifacts that are frequently seen in endogenously caused motion such as PLANE, CAR, and TRAIN. Examples of non-mover concepts include ROCK, CARABINEER, and KNISH. Salient perceptual features from the representational content of the antecedent perceptual state are recorded in one of three feature spaces that correspond to the three global categories. The transition from perceptual representations to a global category feature space is achieved by biases that rely on different perceptual cues to select one of the three concept formation paths. I drew on the developmental psychology literature to motivate the relevant salient perceptual features. For example, biological motion proved to be a particularly salient feature. Biological motion is a perceptual cue that biases the concept formation system to choose the animate feature space. From feature space recordings, representations of feature clusters begin to emerge. With time

and often with the aid of linguistic cues, a new concept is formed and associated with a feature cluster that represents the typical superficial properties of the property indicated by the newly formed concept. After some new concept, C, is formed, a stereotype of its extension can be created and associated with C. The stereotype draws on the information represented in the feature cluster and stored perceptual memories of instances of the property indicated by C. The stereotype along with stored perceptual memories function as heuristics for future use of the concept C.

The GEM concept formation model provides a regular pattern of transitions that begin with perceptual representation and eventuates with a newly formed concept and an associated stereotype. It shows that we can form GEM concepts using very little information, and thus that we can possess concepts despite knowing almost nothing about the properties the newly formed concepts indicate.

After presenting the GEM concept formation model I described a normative model for belief revision that reflects the different stages of concept formation. In the normative model, each stage of the concept formation model has a corresponding defeasible presumption: The defeasible presumptions are as follows:

Presumption 1: “Veridical Perception Presumption”

The perceptual system is reliable in the present environment, and it produces perceptual states that are veridical.

Presumption 2: “No Global Fakes Presumption”

There are no cross-global-type fakes.

Presumption 3: “One-Look-One-Kind Presumption”

There is one kind for every feature cluster. An unfamiliar looking particular is an instance of a new kind, and a familiar looking particular is a particular of a previously observed kind.

Presumption 4: “Fakes-And-Deviants-Are-Rare Presumption”

The sample on the basis of which a stereotype is formed is not made up of fakes or deviants.

All of the defeasible presumptions are underwritten by a presumption of normality, that is, the presumption that the present environment is the normal environment. The normal environment is the environment in which the biases of the perceptual system and the concept formation system are established and explained. The defeasible presumptions ensure that nearly all of our empirical beliefs that employ GEM concepts are empirically revisable. A belief is empirically revisable if there is an epistemically warranted route to the rejection of the belief, given some possible empirical evidence.

The concept formation model, along with its corresponding defeasible presumptions, explains how the projection problem is solved. Concepts whose instances are look-alikes, as in the case of cats and robocats, would produce the same feature clusters and be associated with the same stereotype. The perceptual connection that runs through the perceptual feature clusters to cats in our world, explains why we form the concept CAT. The perceptual connection to instances of robocats in the robocat world explains why we would form the concept ROBOCAT in the robocat world. However, if we were in the robocat world, our concept ROBOCAT would be associated with the global category ANIMATE instead of NON-LIVING MOVER. But this miscategorization can ultimately be revised given some empirical evidence that shows that robocats are not in fact animates, but are actually robots.⁹⁴ As for coextensive but distinct concepts, the GEM concept model can solve the projection problem, but the solution will turn on the case at hand. The GEM concept formation model begins with three global categories. As an infant develops, her concept formation system will form more global categories. So, for example, because of the small number of developmentally early global-level concept formation paths, it is implausible that the concept GRANNY'S FAVORITE ANIMAL would be formed before the concept

⁹⁴ If we lived in a world with both robocats and cats, the concept produced would be a mongrel concept that could be differentiated into two concepts given some empirical evidence revealing that what we previously took to be instances of one kind are actually instances of two look-alike kinds.

CAT. To pose a real challenge to the GEM concept formation model, we would need an instance of distinct co-extensive concepts that are possibly equals in the order of development. Moreover, the two concepts would have to be sensitive to the same perceptual and linguistic cues under all conditions. Satisfaction of this condition is surely very rare.

Resolving the projection problem does not answer what I called “the epistemic challenge.” Reformulated to fit the GEM concept formation model, the epistemic challenge is as follows:

The Epistemic Challenge: Under what conditions is the formation of a GEM concept on the basis of perceptually derived feature clusters warranted?

In Chapter 2 (section 4) I considered and ultimately rejected an evidential and a (merely) reliabilist condition. I argued that to yield warrant, GEM concepts must be formed on the basis of a process that is non-accidentally reliable with respect to the representational functions of our representational capacities. The process yields a kind of warrant I called “concept formation entitlement”. Concept formation entitlement is a kind of entitlement that is a species of the Burgean notion of epistemic warrant. Epistemic warrant is an epistemic good that is achieved by a psychology’s meeting norms or standards that contribute to the fulfillment of the representational function of forming or preserving true beliefs. Entitlement is warrant without reason. It is warrant that derives purely by a psychology’s operating well and reliably with respect to its representational functions.⁹⁵

The entitlement produced by a token of a warranted concept formation process is a prima facie entitlement to employ the newly formed concept in empirical beliefs. The entitlement is especially important in the case of early empirical beliefs about the unknown. Early empirical beliefs (hereafter simply early beliefs) are the first beliefs that employ a newly formed concept.

⁹⁵ In contrast to warrant by justification, if S is warranted by entitlement, S does not need to represent (either personally or sub-personally) the propositional content that characterizes her warrant.

The unknown is any empirical subject matter about which one knows almost nothing. That we can form and thereby come to possess concepts that indicate properties about which we know almost nothing raises the following question: How can our early beliefs about the unknown be warranted?

Concept formation entitlement explains how we can have warranted early beliefs about the unknown. In doing so, concept formation entitlement provides a missing epistemic link in our epistemological theories of the pursuit of empirical knowledge. Recall the example of the child who discovers a tomato. Concept formation entitlement provides the prima facie warrant for her to employ the newly formed concept TOMATO in her first tomato beliefs.

The prima facie entitlement achieved in forming GEM concepts on the basis of perceptual experiences has two sources: (1) the nature and function of our sense-perceptual capacities by which perceptual states and perceptually derived feature clusters are formed; and (2), the contribution that the process makes to the fulfillment of epistemic norms that govern the formation of empirical beliefs and empirical discovery.

A component of concept formation entitlement is our perceptual entitlement to hold our basic perceptual beliefs. Perception provides the causal connection to particulars in the empirical world, and the biases of the concept formation system. The biases shape the representational content of feature clusters, and underlie the defeasible classification of particulars into global categories. An explanation of empirical discovery has to begin with an explanation of how the representational system picks out particulars in perception, and creates ways of representing those particulars as members of a kind.

Biases based on perceptual cues that determine the selection of the concept formation path track truth and deeper structure because of the regularity of the relationship between

perceptually accessible properties and GEM properties. That is why perceptual entitlement is a component of concept formation entitlement. If the perceptual cues did not track real patterns, a reliable perceptual system that regularly produced veridical perceptual states would not be a good guide to deeper structure.

In Chapter 3 I gave an example of epistemic norms that govern the formation of empirical beliefs and empirical discovery: I called these “revisability norms”. The revisability norms are standards for forming and revising beliefs. The GEM concept formation process contributes to the fulfillment of revisability norms by way of the defeasible presumptions. I explained that the defeasible presumptions support the transition from basic perceptual beliefs to beliefs that require the formation and application of new concepts. The presumptions support that transition in at least two ways: First, like their corresponding biases GEM concept formation, the presumptions reliably track real patterns in the world. And second, they guarantee the revisability of our empirical beliefs.

The defeasible presumptions support the formation of warranted early empirical beliefs by securing the possibility of warranted belief revision. If the presumptions that corresponded to each part of the concept formation process were not defeasible, then the concepts produced by the process would carry non-defeasible ontological commitments. The defeasible presumptions provide an escape route for the ontological commitments carried by empirical beliefs, given some empirical evidence. They secure the possibility of long-run reliable empirical belief formation.

Ultimately, the defeasible presumptions are part of the explanation of the possibility of massive empirical dubitability: The ability to consider and bring into doubt nearly all of our

empirical beliefs (though not all at once).⁹⁶ Throwing empirical claims into doubt merely requires conceiving of some piece of empirical evidence that can launch the unraveling of our empirical commitments. The presumptions mark a path for backpedaling from the types of ontological commitments that are characteristic of early empirical beliefs about the unknown.

⁹⁶ The phenomenon of massive dubitability is highlighted by Burge in his reflections on “the intellectual norms which provide the permanent possibility of challenge to any actual practices of individuals or communities that we could envisage.” (1986/2007b, p. 274)

Bibliography

- Arterberry, M. & Bornstein, M. (2001). Three-month-old infants' categorization of animals and vehicles based on static and dynamic attributes. *Journal of Experimental Child Psychology*, 80, 333-346.
- Arterberry, M. & Bornstein, M. (2002). Infant perceptual and conceptual categorization: the roles of static and dynamic stimulus attributes. *Cognition*, 86, 1-24.
- Bertenthal, B. (1993). Infants' perception of biomechanical motions: Intrinsic image and knowledge-based constraints. In B. Bertenthal & C. Granrud (Eds.), *Visual perception and cognition in infancy* (Vol. 21, pp. 175-214). Hillsdale, NJ: Erlbaum.
- Burge, T. (1993). Content preservation. *The Philosophical Review*, 102(4), 457-488.
- Burge, T. (1996). Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society*, 96, 91-116.
- Burge, T. (2003). Perceptual entitlement. *Philosophy and Phenomenological Research*, 67(3), 503-548.
- Burge, T. (2007a). Individualism and the mental. In *Foundations of mind* (pp. 100-150). New York, NY: Oxford University Press. (Original work published 1979)
- Burge, T. (2007b). Intellectual norms and foundations of mind. In *Foundations of mind* (pp. 254-274). New York, NY: Oxford University Press. (Original work published 1986)
- Burge, T. (2007c). Other bodies. In *Foundations of mind* (pp. 82-99). New York, NY: Oxford University Press. (Original work published 1982)
- Burge, T. (2007d). Postscript to Individualism and the Mental. In *Foundations of mind* (pp. 151-182). New York, NY: Oxford University Press.
- Burge, T. (2009). Five theses on *de re* states and attitudes. In J. Almog & P. Leonardi (Eds.), *The philosophy of David Kaplan* (pp. 246-316). Oxford, UK: Oxford University Press.
- Burge, T. (2010a). *Origins of objectivity*. New York, NY: Oxford University Press.
- Burge, T. (2010b). Steps toward origins of propositional thought. *Disputatio*, 4(29), 39-67.
- Burge, T. (2011). Border crossings: Perceptual and post-perceptual object representation. *Behavioral and Brain Sciences*, 34(3), 125.
- Burge, T. (2013). *Cognition through understanding*. New York, NY: Oxford University Press
- Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press.

- Carey, S. (2011). Precis of *The origin of concepts*. *Behavioral and Brain Sciences*, 34(3), 113-162.
- Carey, S. (2014). On learning new primitives in the language of thought: Reply to Rey. *Mind & Language*, 29(2), 133-166.
- Dewar, K., & Xu, F. (2007). Do 9-month-olds expect distinct words to refer to kinds? *Developmental Psychology*, 43, 1227-1238.
- Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. (1981). The present status of the innateness controversy. In *Representations: Philosophical essays on the foundations of cognitive science* (pp. 257-316). Cambridge, MA: MIT Press.
- Fodor, J. (1994). *The elm and the expert*. Cambridge, MA: MIT Press.
- Fodor, J. (1998). *Concepts: Where cognitive science went wrong*. New York, NY: Oxford University Press.
- Fodor, J. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT Press.
- Fodor, J. (2008). *LOT2: The language of thought revisited*. New York, NY: Oxford University Press.
- Fodor, J, Mandler, J., Keil, F., Gopnik, A., Niyogi, S., Rogers, T., ... Snedeker, J. (2005, July). *Solutions to Fodor's puzzle of concept acquisition*. Symposium conducted at a meeting of the Annual Cognitive Science Society, Stressa, Italy. Transcript retrieved from https://software.rc.fas.harvard.edu/lds/wp-content/uploads/2012/04/Niyogi_Snedeker_2005.pdf
- Gallistel, C. R. & King, A. P. (2009). *Memory and the computational brain: Why cognitive science will transform neuroscience*. Malden, MA: Wiley-Blackwell.
- Gao, T., Newman, E. & Scholl, B. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, 59, 154-179
- Gati, I. & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 325-340.
- Gelman, S. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.
- Goldstone, R. L. & Son, J. Y. (2005). Similarity. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 13-36). New York, NY: Cambridge University Press.

- Goldman, A. I. (1976). What is justified belief? In G. Pappas (Ed.), *Justification and knowledge* (pp. 1-23). Dordrecht: D. Reidel.
- Greenberg, M. (2013). *Incomplete understanding, deference, and the content of thought* (UCLA School of Law Research Paper No. 07-30). Retrieved from SSRN Database. (<https://ssrn.com/abstract=1030144>)
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience* 12(5), 711-720.
- Huntley-Fenner, G., Carey, S. & Solimando, A. (2002). Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition*, 85, 203-221.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201-211.
- Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Keil, F. (2008). Innateness. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences*. Retrieved from <http://cognet.mit.edu/library/erefs/mitecs/keil1.html>.
- Kellman, P. J. & Arterberry, M. A. (2006). Infant visual perception. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (6th ed., Vol. 2, pp. 109-160). Hoboken, NJ: John Wiley & Sons.
- Landau, B., Smith, L. & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3, 299-321.
- Landau, B., Smith, L. & Jones, S. (1992a). Perception, ontology, and naming in young children: Commentary on Soja, Carey, and Spelke. *Cognition*, 43, 85-91.
- Landau, B., Smith, L. & Jones, S. (1992b). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language*, 31, 807-825.
- Laurence, S. & Margolis, E. (2002). Radical concept nativism. *Cognition*, 86, 25-55.
- Leslie, A. (1982). The perception of causality in infants. *Perception*, 11, 173-186.
- Leslie, A. & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, 25, 265-288.
- Macnamara, J. (1986). *A border dispute: The place of logic in psychology*. Cambridge, MA: MIT Press.
- Mandler, J. (2004). *The foundations of mind*. New York, NY: Oxford University Press.

- Mandler, J. & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, 8, 291-318.
- Mandler, J. & McDonough, L. (1998). On developing a knowledge base in infancy. *Developmental Psychology*, 34(6), 1274-1288.
- Margolis, E. (1998). How to acquire a concept. *Mind & Language*, 13, 347-369.
- Margolis, E. & Laurence, S. (2011). Learning matters: The role of learning in concept acquisition. *Mind & Language*, 26(5), 507-539.
- Medin, D. (1989). Concepts and conceptual structure. *American Psychologist*, 44(12), 1469-1481.
- Medin, D. & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.) *Similarity and analogical reasoning* (pp. 175-195). Cambridge, UK: Cambridge University Press.
- Medin, D. & Rips, L. (2005). Concepts and categories: memory, meaning, and metaphysics. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 37-72). New York, NY: Cambridge University Press.
- Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Oved, I. (2009). Baptizing meanings for simple concepts. In B. Kim (Ed.), *Proceedings of the Twentieth Midwest Artificial Intelligence and Cognitive Science Conference* (pp. 104-109). Fort Wayne, IN: Indiana University-Purdue University Fort Wayne.
- Pauen, S. & Träuble, B. (2009). How 7-month-olds interpret ambiguous motion events: Category-based reasoning in infancy. *Cognitive Psychology*, 59(3), 275-295.
- Plantinga, A. (1995). Précis of warrant: the current debate and warrant and proper function. *Philosophy and Phenomenological Research*, 55(2), 393-396.
- Prinz, J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: MIT Press.
- Pryor, J. (2005). There is immediate justification. In M. Steup & E. Sosa (Eds.), *Contemporary debates in epistemology* (pp. 181-202). Malden, MA: Blackwell Publishing.
- Pryor, J. (2001). Highlights of recent epistemology. *The British Journal for the Philosophy of Science*, 52(1), 95-124.
- Putnam, H. (1962). It ain't necessarily so. *Journal of Philosophy*, 59(22), 658-671.
- Putnam, H. (1975a). The analytic and the synthetic. In *Mind, language and reality: Philosophical papers* (Vol. 2., pp. 33-69). Cambridge, UK: Cambridge University Press. (Original work published 1962)

- Putnam, H. (1975b). Explanation and reference. In *Mind, language and reality: Philosophical papers* (Vol. 2, pp. 196-214). Cambridge, UK: Cambridge University Press. (Original work published 1973)
- Putnam, H. (1975c). Is semantics possible? In *Mind, language and reality: Philosophical papers* (Vol. 2, pp. 139-152). Cambridge, UK: Cambridge University Press. (Original work published 1970)
- Putnam, H. (1975d). The meaning of 'meaning.' In *Mind, language and reality: Philosophical papers* (Vol. 2, pp. 215-271). Cambridge, UK: Cambridge University Press. (Original work published 1975)
- Rey, G. (2014). Innate *and* learned: Carey, mad dog nativism, and the poverty of stimuli and analogies (yet again). *Mind & Language*, *29*, 109-132.
- Rhodes, M & Gelman, S. (2009). Developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*, 244-274.
- Rhodes, M., Leslie, S. & Tworek, C. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526-13531.
- Rochat, P., Morgan, R. & Carpenter, M. (1997). Young infant's sensitivity to movement information specifying social causality. *Cognitive Development*, *12*, 537-561.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.
- Rosch, E. & Lloyd, B. (Eds.). (1978). *Cognition and categorization*. Hillsdale, NJ: Erlbaum.
- Samet, J. & Flanagan, O. (1989). Innate representations. In S. Silvers (Ed.), *Rerepresentation* (pp. 189-210). New York, Kluwer Academic Publishers.
- Samuels, R. (2002). Nativism in cognitive science. *Mind & Language*, *17*, 233-265.
- Samuels, R. (2007). Is innateness a confused notion? In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Foundations and the future* (pp. 17-36). New York, NY: Oxford University Press.
- Schellenberg, S. (2014). The epistemic force of perceptual experience. *Philosophical Studies*, *170*, 87-100.
- Schlottmann, A. & Surian, L. (1999). Do 9-month olds perceive causation-at-a-distance? *Perception*, *28*, 1105-1113.
- Scholl, B. (2001). Objects and attention: The state of the art. *Cognition*, *80*, 1-46.

- Scholl, B. & Tremoulet, P. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8), 299-309.
- Sellars, W. (1956). Does empirical knowledge have a foundation? In H. Feigl & M. Scriven (Eds.), *The foundations of science and the concepts of psychology and psychoanalysis* (pp. 293-300). Minneapolis: University of Minnesota.
- Siegel, S. (2006). Which properties are represented in perception? In T. Szabo Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481-503). Oxford, UK: Oxford University Press.
- Siegel, S. (2010). *The contents of visual experience*. New York, NY: Oxford University Press.
- Silins, N. (2011). Seeing through the ‘veil of perception.’ *Mind*, 120, 329-367.
- Silins, N. (2012). Explaining perceptual entitlement. *Erkenntnis*, 76(2), 243-261.
- Spelke, E. (1990). Principles of object perception. *Cognitive Science*, 14, 29-56.
- Spelke, E., Kestenbaum, R., Simons, D. & Wein, D. (1995). Spatio-temporal continuity, smoothness of motion, and object identity in infancy. *British Journal of Developmental Psychology*, 13, 113-142.
- Spelke, E. & Newport, E. (1998). Nativism, empiricism, and the development of knowledge. In R. M. Lerner (Ed.), *Handbook of child psychology* (Vol. 1, pp. 275-340). New York: Wiley.
- Sterelny, K. (1989). Fodor’s nativism. *Philosophical Studies*, 55, 119-141.
- Stevens, M. (2000). The essentialist aspect of naive theories. *Cognition*, 74, 149-175.
- Stevens, M. (2012). Theoretical terms without analytic truths. *Philosophical Studies*, 160, 167-190.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327-352.
- Watson, J. (1972). Smiling, cooing, and ‘the game.’ *Merrill-Palmer Quarterly of Behavior and Development*, 18(4), 323-339.
- Waxman, S. & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition*, 95, B59–B68.
- Weber, T. (2016). *How to acquire a concept well (epistemically speaking)*. Unpublished manuscript, Department of Philosophy, University of California, Los Angeles.
- Weiskopf, D. (2008). The origins of concepts. *Philosophical Studies*, 140, 359-384.
- Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, 85, 223-250.

- Xu, F. (2005). Categories, kinds and object individuation in infancy. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time: Papers from the 32nd Carnegie symposium on cognition* (pp. 63-89). Hillsdale, NJ: Earlbaum.
- Xu, F. (2007). Rational statistical inference and cognitive development. In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Foundations and the future* (Vol. 3, pp. 199-215). New York, NY: Oxford University Press.
- Xu, F. & Carey, S. (1996). Infant's metaphysics: The case of numerical identity. *Cognitive Psychology*, 30(2), 11-53.
- Xu, F., Carey, S. & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology*, 49, 155-190.
- Xu, F. & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences*, 105(13), 5012-5015.
- Younger, B. (1985). The segregation of items into categories by ten-month-old infants. *Child Development*, 56(6), 1574-1583.
- Younger, B. (1990). Infant's detection of correlations among feature categories. *Child Development*, 61(3), 614-620.
- Younger, B. & Cohen, L. (1983). Infant perception of correlations among attributes. *Child Development*, 54(4), 858-867.