

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Conceptual Atoms and Conceptual Structures: A Proposed Mediation Between Philosophy and Scientific Psychology on the Nature of Conceptual Representations

Permalink

<https://escholarship.org/uc/item/64c0d5cf>

Author

Britton, Christopher James

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Conceptual Atoms and Conceptual Structures: A Proposed Mediation Between Philosophy
and Scientific Psychology on the Nature of Conceptual Representations

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

By

Christopher James Britton

Committee in charge:

Professor Kevin Falvey, Chair

Professor Aaron Zimmerman

Professor Thomas Barrett

September 2023

The dissertation of Christopher Britton is approved.

Aaron Zimmerman

Thomas Barrett

Kevin Falvey, Committee Chair

September 2023

Conceptual Atoms and Conceptual Structures: A Proposed Mediation Between Philosophy
and Scientific Psychology on the Nature of Conceptual Representations

Copyright © 2023

By

Christopher Britton

ACKNOWLEDGEMENTS

I would like to start by extending my thanks to everyone who has helped me over the years of my graduate program, whatever form that help may have taken. I would never have gotten to this point without all of you. While I have attempted to be as comprehensive as possible with my acknowledgements, my poor memory has likely ensured that someone has been left out. I apologize to anyone I may have here neglected.

First and foremost, thank you to my committee members Kevin Falvey, Aaron Zimmerman, and Thomas Barrett. To Kevin, without your courses and seminars, and the many hours of struggling through the most difficult philosophical texts I have ever engaged with, I would never have arrived at the philosophical views which form the foundation for this dissertation. Thank you for all your kindness and patience during those times when life got in the way of my academic work. To Aaron, thank you the many courses and seminars you taught on the philosophy of mind, your support for my work, and for helping me with the term paper that ultimately formed the basis for this dissertation. And thank you for encouraging me during those times when I struggled with my work. To Thomas, thank you for your thorough feedback and meetings on this dissertation, your consistently useful advice on navigating the world of academia, and for serving as an inspiration and model for my own teaching.

Second, I want to express my thanks to all of my fellow graduate students who provided invaluable conversation and feedback throughout the writing of this dissertation. Special thanks to Jason Hanschmann, Daniel Story, and Arnel Blake Batoon, whose many late night conversations helped me see where gaps existed in my arguments and who helped me immensely with clarifying and solidifying my ideas. Thank you as well to Jeff Bagwell, David

King, Tim Mainwaring, Kyle Dickey, Jordan Neidlinger, Jenna Schaal-O'Connor, Sami Garcia, Seyed Yarandi, Blake Kyler, Sam Zahn, Damien Barnes, Thainá Coltro Demartini, Tom Costigan, and Sherri Conklin. Our many conversations both within and without the halls of our department and university expanded my knowledge of philosophy in ways I could never have accomplished without you all.

Lastly, thank you to all the family and friends who provided their company, conversation, and emotional support throughout what were often difficult years. Thank you to my parents, Shari Cook and John Britton, for their continual love, care, and assistance throughout these many years. Thank you to my sister Anna Britton for our many conversations and the good humor and enjoyment they brought. And thank you to my friends Jian (Frank) Wang, Jarrod Salazar, Alec Aragon, Lucia Reynosa, and Miguel Alcantara for providing much needed respite from my academic work, and for providing me with some of my happiest memories during my time here in Santa Barbara.

VITA OF CHRISTOPHER BRITTON

September 2023

EDUCATION

Doctor of Philosophy in Philosophy, University of California, Santa Barbara, September 2023
(expected)

Master of Arts in Philosophy, University of California, Santa Barbara, September 2022

Bachelor of Arts in Philosophy, Western Washington University, June 2014

Associate of Arts, Columbia Basin College, June 2012

PROFESSIONAL EMPLOYMENT

2019-2023: Teaching Associate: Department of Philosophy, University of California, Santa Barbara

2016-2022: Teaching Assistant: Department of Philosophy, University of California, Santa Barbara

2022-2023: Teaching Assistant: Department of Writing, University of California, Santa Barbara

2018: Course Reader: Department of Philosophy, University of California, Santa Barbara

2013-2014: Course Reader: Department of Philosophy, Western Washington University

AWARDS

Paul Wienpahl Award for Excellence in Teaching, 2020-2021

R.W. Church Scholarship, UC Santa Barbara, 2016-2022

Departmental Fellowship, UC Santa Barbara, 2015-2016

FIELDS OF STUDY

Areas of Specialization: Philosophy of Mind, Philosophy of Cognitive Science

Areas of Competence: Political Philosophy

ABSTRACT

Conceptual Atoms and Conceptual Structures: A Proposed Mediation Between Philosophy
and Scientific Psychology on the Nature of Conceptual Representations

By

Christopher Britton

My dissertation concerns the relation between two ways of understanding conceptual representations in cognitive science, that had by philosophers and that had by psychologists. For philosophers, concepts are supposed to make possible the representation of the world in thought and to act as building blocks for those representations which support our having of propositional attitudes, such as belief and desire. Psychologists, by contrast, see concepts as complex mental representations that function to guide practical activities such as categorizing and reasoning about the world around us.

Typically it has been thought by both philosophers and psychologists that both sides are concerned with the same thing when they use the word “concept,” and much has been said about how to reconcile the two approaches. In my dissertation I argue that such a reconciliation is not forthcoming, owing to the different and in tension functions concepts have been expected to perform by each side of the debate.

To remedy this, I argue that a distinction ought to be made between what I term conceptual atoms and conceptual structures. It is conceptual atoms that I take many

philosophers to be referring to when they use the word “concept,” while psychologists instead refer to conceptual structures. Conceptual atoms are semantically unstructured mental representations which combine together to form the aforementioned conceptual structures, with these structures forming an open-ended class of mental representation that includes the prototypes, exemplars, and folk theories posited by psychology. It is these conceptual structures that both guide our practical activities and support our having of propositional attitudes. In short, what are called “concepts” by philosophers are the building blocks of what are called the same by psychologists.

In chapter 1 I lay the groundwork for the dissertation as a whole by surveying some of the foundational issues for the study of concepts in both philosophy and psychology, including the explanatory roles for concepts as well as issues surrounding how to understand conceptual individuation and structure. In chapters 2-3 I survey the leading theories of concepts in psychology and argue that a pluralist theory allowing for a wide variety of structured concepts is most plausible given the available evidence. In chapter 4 I investigate the main philosophical competitor to psychological theories of concepts, namely conceptual atomism, and argue for a particular interpretation of it. In chapter 5 I argue for my own alternative view that divides concepts into conceptual atoms and structures and consider some of its implications and some of the difficulties that it must confront. In chapter 6 I argue that the study of conceptual atoms and structures should adopt an externalist and ecological approach that studies conceptual structures as they are actually used by normal thinkers, in environments normal to them, and in service of the ends they normally seek. Finally, in chapter 7 I argue that teleosemantics is best equipped to provide the form of explanation and theory of representation required by such an approach.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
VITA.....	vi
ABSTRACT.....	vii
Introduction	1
1 – What (If Anything) Are Concepts?	8
1.1 – How to Approach the Questions.....	11
1.2 – A Preliminary Answer	17
1.3 – Two Oft-Competing Roles	21
1.4 – Concept Individuation	34
1.5 – Conceptual Structure	42
2 – Theories of Concepts.....	49
2.1 – The Classical Theory	52
2.2 – The Prototype and Exemplar Theories	65
2.3 – The Theory-Theory	73
2.4 – Other Theories	83
2.4.1 – Neoclassical Concepts	83
2.4.2 – Ideals.....	85
2.4.3 – Ad hoc Concepts.....	86

2.4.4 – Hybrid Theories and Concept Pluralism.....	89
3 – Conceptual Pluralism	92
3.1 – Coarse Conceptual Pluralism.....	92
3.2 – Fine Conceptual Pluralism.....	98
3.2.1 – Differences in Experience	100
3.2.2 – Differences in Stage of Cognitive Development.....	102
3.2.3 – Differences in Capacities	103
3.2.4 – Differences in Context	104
3.2.5 – Differences in Purpose.....	105
3.3 – Objections.....	108
3.3.1 – Background and Conceptual Knowledge.....	108
3.3.2 – Hybrid Theories of Concepts	119
4 – Concept Individuation: Atomism Reconsidered	124
4.1 – What is Molecularism?.....	126
4.2 – The Appeal of Molecularism in Psychology	127
4.3 – The Appeal of Molecularism in Philosophy.....	129
4.3.1 – Molecularism and Frege Cases	130
4.3.2 – Molecularism and Normative Demands on Concept Possession.....	133
4.4 – What is Atomism?.....	140
4.4.1 – Proposal CA1: Concepts Lack Structure	141

4.4.2 – Proposal CA2: Concepts Lack Structure Relevant to Psychology.....	143
4.4.3 – Proposal CA3: Concepts Lack Representational Structure.....	144
4.4.4 – Proposal CA4: Concepts Lack Individuative Representational Structure	150
4.4.5 – Proposal CA5: Concepts Are Referentially Individuated at a High-Level of Abstraction and Multiply Realizable	155
4.5 – Some Potential problems	165
5 – From Concepts to Conceptual Atoms and Structures.....	168
5.1 – Conceptual Atoms and Conceptual Structures.....	171
5.2 – Two Arguments Against Traditional Concepts.....	174
5.2.1 – The Unneeded Proliferation of Concept Kinds.....	177
5.2.2 – The Epistemic Theory of Natural Kinds	179
5.2.3 – Natural Fine Structural Kinds	187
5.2.4 – The Functional Role of Concepts.....	190
5.2.5 – Some Objections.....	194
5.3– Conceptual Atoms and Conceptual Structures	203
5.3.1 – Some Outstanding Issues.....	211
5.4 – Concept Eliminativism.....	214
5.5 – Takeaways and Potential Lessons	220
6 – The Need for an Externalist Semantics for Conceptual Atoms and Structures.....	225
6.1 – From Concepts to Conceptual Atoms and Structures	225

6.2 – The Need for an Externalist Semantics for Conceptual Representations.....	232
6.3 – Externalist Semantics and Conceptual Variability.....	237
7 – Towards a Teleosemantics for Conceptual Atoms and Structures.....	244
7.1 – Externalist and Teleosemantic Explanations.....	245
7.2 – Teleofunctions for Conceptual Atoms and Structures	251
7.3 – A Teleosemantic Approach for Conceptual Atoms and Structures	262
7.3.1 – Exploitable Relations.....	264
7.3.2 – Exploitable Causal Correlations.....	265
7.3.3 – Exploitable Structural Correspondence.....	267
7.3.4 – Normal Explanations for the Exercise of Core Cognitive Capacities.....	271
7.4 – Advantages to the Approach	276
7.5 – Some Outstanding Issues	282
Conclusion.....	286
Works Cited	291

Introduction

Talk of concepts is ubiquitous in both academic and daily life. Despite this, they remain highly obscure entities within both domains. This is due, in part, to the myriad of ways in which the word “concept” is used, as well as the wide variety of roles that concepts, whatever they turn out to be, have been expected to fill. In the broadest of terms, my dissertation is focused on what concepts must be like in order to do all that has been expected of them. In doing so, I hope to mediate between philosophers and scientific psychologists in particular, who are often at odds due to their sometimes starkly different ways of thinking about concepts.

Arguably the most foundational role that concepts play is allowing us to learn about the world around us and to apply that knowledge. Nearly every task we perform in day-to-day life requires the reidentification of things we have previously encountered and the acquisition and application of knowledge regarding them. In baking a cake I must be able to identify which crystalline substance in my cabinets is sugar and which is salt, and to apply my knowledge of which of these two is called for by the recipe and when. Such tasks are performed so effortlessly that we are liable not to appreciate the immense cognitive complexity making their execution possible. It is concepts that have long been understood as the mental representations which underly core cognitive capacities such as these that are so foundational to life. It is only because I have concepts of sugar, salt, cakes, ovens, and so on, that I can even hope to carry out the task of baking a cake in the first place.

This is not all that concepts do for us. They have another important theoretical role, though unfortunately one which is in tension with their role as cognitive tools for reidentification and knowledge acquisition/application. Concepts are supposed to be the building blocks of thought,

those things one must possess in order to form thoughts and to thus think about the various aspects of one's environment. One's beliefs, desires, hopes, and dreams are built out of one's concepts, perhaps in the same way that the sentences one speaks and writes are built up out of words and phrases. Thus, forming such propositional attitudes requires that one possess the concepts from which they are built. I cannot believe that there is leftover cake in the fridge unless I have concepts of cakes and fridges, nor can I desire to eat it, hope there is some left, fear my roommate ate it, worry I have consumed too much, and so on.

My dissertation is broadly focused on how to best understand concepts and the roles they play in our mental lives. Traditionally, it has been thought that possessing a concept requires that one have the right set of beliefs or be disposed to draw the right inferences. In other words, it has long been thought that concepts are essentially normative, that possessing one requires that one meet a certain set of normative demands. On this way of thinking, in order for me to possess the concept of sugar, I must believe or be disposed to infer that sugar is a white crystalline substance, that it is sweet, that it feels rough when rubbed between two fingers, and so on. Were I instead to think the substance in my hand is a reddish, slimy, sour substance, I would *ipso facto* not possess a concept of sugar on account of my confusion and ignorance. Or perhaps I must be able to recognize sugar when I encounter it, sort sugar from non-sugar, and distinguish sugar from substances with which it is superficially similar. Taking this traditional route forges a tight and inseparable connection between (1) the adequacy of a concept for various tasks and (2) the capacity to think about the relevant aspects of one's environment. In virtue of this connection, one cannot so much as think about something unless one possesses a cognitive tool adequate for engaging with it productively.

This has two problematic results. First, those with beliefs different from the norm or those inclined to draw nonstandard inferences are foreclosed from even thinking of the relevant aspects of their environments. The two groups most often denied concepts on this basis are young children and nonhuman animals, leading philosophers to an impoverished understanding of their mental lives and a lack of appreciation for its true complexity. However, even with adult humans, the traditional approach fails to accommodate the immense diversity present in human thought and the wide variety of tasks concepts are deployed in service of. The second problem is that it rules out the possibility of legitimate conceptual development. If adequacy is demanded as a condition for possessing a concept in the first place, there is no way to give a plausible theory as to how thinkers grow more adept at using concepts over time to achieve their various aims.

I ultimately argue for a more pluralistic and ecologically minded approach to understanding conceptual representations, one which avoids these two problems. There is no one class of representation called “concepts.” Rather, there are conceptual atoms on the one hand and conceptual structures on the other. The former is what the philosopher has in mind when they speak of concepts as the building blocks of thought, while the latter are what psychologists have in mind when they talk of concepts as cognitive tools for guiding behavior in service of one’s aims. Atoms serve as the building blocks for conceptual structures, and it is these structures which both guide the exercise of our cognitive capacities and support the attitudes like belief and desire that philosophers have long concerned themselves with.

So conceptual structures are, in part, tools for thinking about and productively engaging with the world. But importantly there is no one task any given conceptual structure is called upon to perform, nor is there any one set of beliefs or inferential dispositions required to

possess a given conceptual structure. The approach is thus pluralistic in that I advocate a view under which conceptual structures can take many forms and, relatedly, that there are many paths to possessing any given conceptual structure and its component atoms. In doing so, I make space for greater variability in ways of thinking than is allowed on more traditional approaches. Further, the approach is ecologically minded in that I argue that any adequate theory of concepts must attend to facts such as what the thinker is trying to do, what environment they find themselves in, as well as the capacities and information available to them in trying to make their way in life. Only by attending to such facts can one hope to discern what form a given thinker's conceptual structures take and when/how they are adequate.

Perhaps the guiding goal of all that follows here is that of placing concepts in an ordinary, lived-in setting. I am interested in how concepts are actually used in real-world contexts in the service of satisfying the user's actual needs and desires. There is a tendency, I believe, for philosophers to want concepts to be insulated from the everyday messiness and instability of human thought and activity. This has led to a common picture of concepts that, far from illuminating the nature of human thought, serves to obscure it. Only by looking at how concepts are actually used, as well as the myriad aims their users put them towards, can concepts truly be understood.

The main justification for the view I defend here is not any arguments directly against the traditional view's possibility. Ultimately, I do not think that there is anything incoherent in the traditional picture of concepts. Further, given a suitable set of explanatory goals and assumptions, the traditional picture may even be reasonable. However, it is these goals and assumptions that are, I believe, misguided. What I attempt here is to articulate an alternative set of goals and assumptions, and with them a picture of concepts, one with enough flexibility

and with limited enough demands that it is able to accommodate the full complexity and variety of cognition across all thinking beings.

As with most topics in philosophy and elsewhere, discussions of concepts lack a clean and obvious entry point. Understanding one facet of concepts requires first understanding many others, with the same true of these other facets as well. As such, much of the early dissertation will be focused on laying the groundwork for the more original work that comes later.

Chapter one is focused on a variety of foundational questions regarding the nature of concepts and their study. It covers first primarily methodological questions, such as how to approach the study of concepts as well as what their theoretical roles in both philosophy and psychology have been taken to be. After this, I will cover what is likely the most discussed question in the philosophy of concepts, namely how concepts are individuated, as well as the closely related question of whether, and if so how, they are structured.

Chapter two surveys each of the major theories of concepts. Of primary focus will be the main theories proposed by cognitive and developmental psychologists, namely the prototype theory, exemplar theory, and theory-theory. Later in the chapter less conventional theories will be briefly discussed, namely the neoclassical and hybrid theories of concepts. Finally, I look at ad hoc concepts, those concepts which are held to be assembled in short-term memory and discarded after use, rather than stored in long-term memory.

Chapter three argues that none of the major theories of concepts surveyed in chapter 2 is likely to be explanatorily adequate on their own. Due to this, we ought to instead adopt a pluralist theory of concepts, one which holds that concepts can come in wide variety of structured forms. In addition, I look at different forms of concept pluralism one may adopt as

well as some of the reasons that concepts of the same referential type may come to differ in their structures. I end the chapter by arguing that the hybrid theory of concepts is not, as is sometimes thought, a competitor to pluralism, but is rather compatible with it.

Chapter four discusses the debate between molecularism and atomism. In it I argue that atomism as usually stated is obscure, with its central claim being ambiguous. I consider various ways of formulating atomism, ultimately arriving at atomism as the view that concepts are referentially individuated at a high level of abstraction, but multiply realizable by a large number of more concrete concept kinds individuated in terms of their structures. I end the chapter with a consideration of some of the problems my interpretation of atomism faces, both as a view in its own right and qua interpretation.

Chapter five presents my own view. In it I argue that we ought to abandon concepts understood as representations that perform the various roles both philosophers and psychologists have expected of them. In their place I posit what I call conceptual atoms and conceptual structures. Conceptual atoms are, roughly, semantically unstructured and referentially individuated conceptual representations. Conceptual structures are, also rather roughly, concepts as understood in psychology, e.g., prototypes, exemplars, theories, etc. On my view, conceptual atoms are the constituents of conceptual structures, and it is the structure formed by conceptual atoms when combined that encodes the information in concepts that is exploited in action. I end the chapter with an argument against concept eliminativism and a look at some potential applications for my view.

Chapter 6 argues that the position laid out in chapter 5 is more of a framework for a theory of conceptual representations rather than a theory proper. This is because it leaves a major gap unfilled: how conceptual atoms and structures acquire their referents. I survey proposals from

psychology and internalist approaches for how concepts (and in my case, conceptual atoms) may acquire their referents. I argue that these proposals all either impose unmeetable conditions on ordinary conceptual atom possessors or deliver the wrong referents for both conceptual atoms and structures.

Finally, chapter 7 argues that the approach to semantics that is most likely to satisfy the desiderata for a theory of conceptual atoms and structures is one that is broadly teleosemantic in nature. Drawing from the philosophical work of Ruth Millikan, I argue that conceptual structures have the teleofunction of assisting directive representations in the performance of their teleofunctions. Similarly, conceptual atoms have the teleofunction of combining into semantically significant relations with one another to form conceptual structures that are of use to their possessor. I then end the chapter by looking at some potential applications of the view, most notably how it can help to provide an account of knowledge acquisition of a certain kind.

1 – What (If Anything) Are Concepts?

Concepts are one of the most central and longest studied phenomena in both philosophy and psychology. In philosophy, the study of concepts (as many contemporarily understand them) goes back to at least Hume and Locke, and possibly earlier.¹ In psychology, studies on concepts date back to at least the 1920s.² This longstanding and continued focus on concepts is well-motivated and quite understandable. Both philosophers and psychologists hope that concepts, whatever they turn out to be, will prove able to carry a quite heavy theoretical burden. They are expected to explain cognitive capacities of such importance and centrality, such as the capacity to learn about things and employ that knowledge on later occasions in service of practical ends, that all but the simplest of minds fail to exhibit them.³ Despite their long prominence in both psychological and philosophical theories of the mind, however, surprisingly there has yet to emerge much in the way of consensus regarding concepts. Further, what little progress has been made is largely negative in character. Most work on concepts, especially in philosophy, tells us a good deal about what concepts cannot be if they are to be suited to various explanatory ends. Comparatively little success, however, has been had when it comes to providing a positive story, one telling us what exactly concepts are.⁴ Thus, it

¹ See Fodor (2003) for an argument that Hume's work in particular forms the foundation for much of the current work on concepts in the cognitive sciences, in particular his claims that ideas are mental particulars and that mental processes are only sensitive to the formal features of the ideas that are involved in them. The earliest investigation of something like concepts is arguably Plato's *Euthyphro*, where the concept of piety is considered.

² Hull (1920).

³ This is, admittedly, a much more contentious claim among philosophers than psychologists. I will later argue that, absent certain molecularist or internalist assumptions, there is little reason to refrain from attributing even quite sophisticated concepts to non-human animals and young children.

⁴ For instance, the Classical Theory of concepts, which held that concepts are mentally represented definitions, has since the 70s been largely viewed as a theoretical non-starter. In

remains to be asked: What exactly are concepts?

Despite initial appearances to the contrary, this is hardly a straightforward question. There are a great many issues involved, issues largely independent from one another. These issues can, however, be separated into two broad groups. First, there are issues surrounding the metaphysics of concepts, which primarily take the form of what broad ontological category, e.g. object, event, property, etc., concepts should be classed under. Second are issues concerning the proper theoretical role of concepts, i.e. what sort of phenomena we postulate concepts and develop their corresponding theories in order to explain. These two issues are, perhaps surprisingly, treated as largely independent from one another in the actual practice of many philosophers and psychologists. As we will soon see, theorists who differ quite radically on the metaphysical issues often agree on the issues concerning theoretical roles, and vice versa.⁵

Given all this, we can separate our initial question into two, narrower ones:

1. What kind of thing are concepts, i.e. what ontological category do concepts belong to?
2. What do concepts do, or what is it that we posit concepts in order to explain?

psychology this was due to experimental results that it failed to account for, e.g. typicality effects. Within philosophy the Classical Theory's influence waned primarily due to externalist arguments against descriptivist approaches to mental content, worries about the viability of the analytic/synthetic distinction, and the seeming unavailability of definitions for many, if not most, commonly encountered phenomena.

⁵ Fodor and many of his opponents in psychology, for instance, agree that concepts are mental representations, even while disagreeing wholly on issues such as concept structure and reference determination. Similarly, those in philosophy and psychology that hold concepts are posited to explain things such as categorization behavior have held that concepts are everything from abilities to mental representations to abstract Fregean senses.

These questions, while distinguishable, are often erroneously treated as if they are independent from one another. Concepts are, as of yet, unobservable. Our only guide to their nature (and even their possible existence) is through their effects on things more readily observed. The study of concepts is then, in a way, akin to the study of a hitherto unknown animal through its spoor.⁶ It is a messy, lurching process, prone to error and revisitation of starting assumptions. It is rather unlike the more reliable and predictable study of the animal via direct observation or through the poking and prodding of it via various scientific instruments. And just as the study of spoor can lead sometimes to fantastical claims as to their cause (think of cryptids and how much can be built on a foundation as flimsy as a few footprint casts), the study of concepts has sometimes resulted in rather strange and often conflicting claims.⁷ As such, we can and should anticipate that much of the discussion to follow will inherit this messiness. Further, much of the work to follow will involve clearing away the accumulation of assumptions and supposed established truths that have generated so much of the long ongoing disputes in the literature.

The remainder of this chapter will focus specifically on the two questions above and the theoretically prior question of how to go about answering them both. The structure of the discussion to come is as follows. First, I will present an admittedly rough, though initially useful, characterization of concepts as ways of thinking. This is intended to provide a sufficient

⁶ It is similar in this respect to any other scientific research into unobservables, such as the various particles posited by physicists.

⁷ These often are a result of the ever-expanding theoretical role for concepts. As will be seen, what concepts are expected to explain by philosophers and cognitive scientists has expanded to the point that it is no longer clear that any one thing could ever hope to explain them all. Further, the things concepts are expected to explain are sometimes so disparate that it is unclear why anyone would think that they are explainable by any single thing, other than simply theoretical baggage accumulated by theories of concepts over the decades. These issues will be discussed thoroughly in ch. 4.

base from which a more elaborate and determinate understanding of concepts can be constructed. Second, I will look at various ways of approaching the question of what concepts are, settling on an approach that understands them as theoretical posits. This approach will focus attention upon the explanatory role played by concepts in the theories of cognitive science. Third, I will look at the two most central of these roles in both philosophy and psychology. Fourth, I will look at the question of concept individuation, in particular getting straight on what exactly the question is concerned with and whether it is really one question at all. Finally, I will examine the notion of conceptual structure, as well as its relation both to issues of concept individuation and to the explanatory roles of concepts.

1.1 – How to Approach the Questions

The questions of what concepts are and what it is that they do are, as previously mentioned, intertwined. In what follows it will be seen that an answer to the second question provides, if not a complete answer, then the beginnings of an answer to the first.

When considering the question of what a concept is, there are at least three approaches to answering it. First, one could attempt to determine what concepts are by examining the use of “concept” and closely related words in ordinary, everyday linguistic contexts. Second, one could look, not at the usage of “concept” among laypeople, but at the usage of the word within “expert” populations, such as in the fields of philosophy and psychology. Finally, one could focus not upon linguistic usage (whether lay or expert) at all, and instead treat “concept” as a theoretical term, with concepts themselves thought of as posits made by the relevant theories. On this approach we consider not how “concept” is used, but on what the entities supposedly referred to by the word must be like in order for the theories positing them to accomplish their explanatory goals. In the following section, I will examine each of these three approaches in

order and ultimately argue for the third, i.e. for treating concepts as theoretical posits in the cognitive sciences.

The first of these ways, i.e. linguistic analysis of ordinary usage, is unlikely to be productive. As amusingly noted by Peacocke, the widespread variability in the use of the word “concept” renders problematic any attempt to answer the question of what concepts are via an examination of our common ways of speaking:

The term “concept” has by now come to be something of a term of art. The word does not have in English a unique sense that is theoretically important. Those who think it does should consider the contribution made to our subject by Woody Allen, when in his film *Annie Hall* he has one of his characters in the entertainment industry say, “Right now it’s only a notion, but I think I can get money to make it into a concept, and later turn it into an idea.” This quotation suggests that we would not make progress by undertaking an analysis of the word “concept” in the style of 1950s linguistic philosophy.⁸

While this is a position I am sympathetic towards, it would I think be instructive to linger on this issue for longer than Peacocke does. While the quote does illustrate the often (to philosophers) odd ways that “concept” gets used, it is only one of the many various ways in which that word is employed in everyday discourse. By considering a few others, we can better appreciate why taking a linguistic analysis approach to our question would be misguided.

In some cases, the word “concept” seems to be used in evaluating one’s competency or capacity to engage in a particular task. Little Timmy grasps the concept MULTIPLICATION when

⁸ Peacocke (1992), pg. 1-2.

he can fill out his times tables with the proficiency expected of a third grader.⁹ But if little Timmy ten years later displays no better proficiency, that attribution may be withheld. In this latter case, Timmy's competency no longer meets the standards expected of someone in his situation. In other cases, the relevant competencies are simpler and less contextually variable. Such attributions vary, not just over time, but space as well. Someone can rightfully be said to grasp some concept (in this sense) in an ordinary setting and not in another, e.g. a professor's grasp of a concept in the authoritative setting of the classroom may be relatively tight, while their grasp loosens substantially in the seminar room or at the conference lectern.

This seemingly tight connection between competencies and concept attributions is best seen with the concepts of basic properties. Some common concepts seem to demand for their possession (or sometimes are even identified with) simple sorting or categorization capacities. Ordinary people are likely to withhold attribution of GREEN and related color concepts to someone who regularly fails to identify and sort objects according to that color. Likewise, few would likely attribute the concept of SQUARE to someone who regularly classed circles and triangles under it.¹⁰

Competencies are not, however, the only thing that is of concern whenever "concept" is used in ordinary contexts. Other times "concept" seems to refer to a culturally and temporally bound understanding of some phenomenon, as when we say things like "Those living in the

⁹ See Wilson (2006) for a defense of the view that concepts are primarily involved in our attributions to others of competency in some task of interest to the attributor (see pg. 23-26 for one instructive example involving the concept RAINBOW). See Millikan (2000), especially ch. 1, 3, and 4, for a view which also understands concepts as abilities (specifically abilities to reidentify substances through various media and in a variety of contexts), but without a focus on attribution.

¹⁰ I will, following common usage, refer to concepts by using small caps, e.g. the concept expressed by the word "snail," and which refers to snails will be itself referred to using SNAIL.

precolonial era had no concept of race” or when those engaged in conceptual engineering advocate for the alteration of a concept in circulation within some community. Here concepts are understood as akin to something like a way of thinking, common in a community at a time, concerning some aspect of the world. Though importantly, on this use concepts need not be accurate or even about anything actually existent in the first place. Medieval Europeans had a concept WITCH that they confidently applied, despite there being no such things. Similarly, concepts of things such as racial and ethnic groups are likely to reflect the prejudices and ignorance of the communities in which they’re found, rather than any objective reality.

Finally, “concept” sometimes is used in the way most common to philosophers, referring to an abstract idea which transcends any one time, place, or culture, as when we say, “There are concepts so complex that no human being could ever grasp them.” Here, unlike with the more bounded understanding of concepts as communal ways of thinking, concepts are closer to *ideal* ways of thinking. The ways of thinking are ideal in that they are both true and get at the “deep” properties of the thing in question, usually by stating substantive necessary and sufficient conditions for membership in the kind corresponding to the concept. This understanding of concepts and this focus on ideals is one that will receive significant attention later in chapter 5.

These latter two uses of “concept” reveal a rather large gap between two common uses of the word. Specifically, there is a large gap between concepts understood as public and as private entities. We often do talk about *the* concept of something, such as or “The concept of numerical identity is surprisingly difficult for undergraduate students to grasp.” Insofar as we speak in this way, we seem to understand concepts as the public, immutable, ideals that philosophers often seem to have in mind. But just as often we talk about a particular person or

group's concept of something and compare it to the concepts of others. We can say "The concept of justice for the Greeks was very different than that of contemporary Americans" and be readily understood as saying that the Greeks thought about justice in different ways than contemporary Americans do. This second way maps cleanly onto the understanding of concepts as communal ways of thinking. The first way of talking of concepts implies something public and immutable, while the second implies something more private and prone to variation across space and time.

Each of the above uses of "concept" corresponds to a general way of understanding what concepts are. Depending on which use one focuses one's attention on, concepts seem to be anything from competencies or abilities, to communal or ideal ways of thinking. However, there are other uses of the word that do not clearly indicate any one understanding of what concepts are. These latter uses are, in a sense, neutral on many questions concerning the nature of concepts. They do, however, shed some light on the multifarious and conflicting ways in which "concept" is used.

In many cases, "concept" is used to talk about some aspect of the world. However, oftentimes "concept" and its variants are used as a modifier for concrete entities in the real world. We talk about works of art being "conceptual" when they are sufficiently avant-garde or experimental. In music, a concept album is simply an album with an overarching theme or with greater cohesion than a typical album. Concept art, often used in the development stages for television, movies, and video games, depicts an artist's interpretation of what various elements of the finished work may look like, serving as inspiration for others working on the project. Concept cars are experimental cars, designed to show off features that are not yet (but are intended to soon be) available to the general consumer. And so on. In each of these cases,

the word “concept” seems to do little more than to indicate that the thing in question is more experimental or, in a rather vague sense, “headier” or more “intellectual” than things similar to it. Or perhaps to point towards something not yet, but hoped to soon be, existent.

The lesson here seems to be that there is little commonality between our various uses of “concept” in English. About the only thing that the word “concept” can be said to do consistently, it seems, is to indicate that one is talking about something more abstract or in a vague sense more “mental” than what the concept is of. Shifting from talk involving “beauty” to talk involving “the concept of beauty” seems to often do little more than (1) shift the discussion from beauty itself to how a person or group or people think about beauty, or (2) shift the discussion from beauty as realized in a particular beautiful object to abstract beauty or beauty-in-itself. Given the wide variation in how “concept” is used and how often these uses come into conflict, linguistic analysis seems unlikely to be a productive approach to understanding concepts.

Even if our use of “concept” were more consistent, it is not clear that our common usage of the word and those related to it would or should be of much interest for our purposes. Biologists do not care much about the pre-theoretic usage of species terms, with the lay tendency to categorize whales as fish rather than mammals left as errata for psychologists to study rather than data biologists must accommodate.¹¹ Nor is the tendency for people to refuse to categorize lemonade and soda as water something that chemists work to understand.¹² It is unclear why the shift to concepts and thus to psychology should change such matters.

¹¹ While this is a common example among both philosophers and psychologists, I myself cannot claim to have ever seen anyone make this particular mistake in the wild. Similar mistakes are, however, common in daily life.

¹² Malt (1994).

Rather than engaging in something akin to 1950s linguistic analysis then, the typical approach in the philosophical literature (at least amongst philosophers with broadly naturalistic inclinations) has been to think of concepts as theoretical entities, ones postulated by theorists to explain certain phenomena under investigation. Among psychologists, this assumption is so longstanding and so universally held, that it seems to fade entirely into the background and escape explicit mention.¹³ By focusing on the phenomena which concepts are marshalled to explain, one can understand what concepts are via the properties they must have in order for one's theory to be explanatorily adequate.

1.2 – A Preliminary Answer

With an approach to the question tentatively decided on, we can now shift to developing a working picture of what concepts are and what they do. By the nature of the approach, there is no ideal entry point to this process. The picture of concepts developed emerges in a way more akin to sculpture than writing, with gradual refinements of form until the whole comes into focus. With that said, we will start in what is as good a place as any: with the question of what broad ontological category concepts are most likely to fall under.

As expected, there is wide variation in opinion on this matter. Theories on the metaphysics of concepts have variously held that they are mere abstracta (Peacocke 1992)¹⁴, abilities (Dummett 1993, Millikan 2000), mental particulars (Fodor 1995, Laurence and Margolis 2007, Millikan 2017), or some combination of the above (Laurence and Margolis 2007).

¹³ It is not even clear that the methodology of psychologists studying concepts would be explicable were they not to hold this assumption. Further, even if “concept” had a determinate meaning in ordinary language, that would not preclude the use of a phonologically identical word in psychology to refer to a theoretical posit.

¹⁴ “Mere” abstracta is meant to distinguish such positions from those that hold that there are abstract concept types corresponding to token abilities, mental representations, and so forth. A mere abstract entity lacks any concrete instance or realization.

Disagreement on this particular issue has occurred largely between philosophers. Psychologists from the 70s onward have tended near universally, when they commit to a position at all, to hold that concepts are mental representations which factor in the processes underlying cognitive capacities like categorization and linguistic comprehension.¹⁵ To what degree these debates are substantive and to what degree the parties involved are simply theorizing about different phenomena is often unclear, such as the disputes over whether psychologists are really studying concepts or rather studying mere conceptions.¹⁶

In what follows, I will assume that concepts are mental particulars. More specifically, I will assume that they are mental particulars of the sort posited by the Computational Theory of Mind (CTM), i.e. mental symbols with both semantic and syntactic properties, computationally processed in accordance with syntactic properties in a way which respects their semantic properties.¹⁷ My reasons for doing this are threefold. One, something akin to this view seems to be assumed by most of those working within the psychology of concepts. For instance, the psychologist Lloyd Komatsu describes a concept as “the mental

¹⁵ See Machery (2009, 2010). I omit Machery’s requirement that such representations be those used “by default”, i.e. that they are representations which can be distinguished from background beliefs and are privileged insofar as they are deployed consistently when subjects are under time pressures. See Malt (2010) and Rey (2009) for arguments that this additional requirement is unwise. I further omit Machery’s requirement that concepts be stored in long-term memory to avoid ruling out the possibility of ad hoc concepts, i.e. concepts assembled on the fly to guide behavior. See Barsalou (1983) for arguments for why at least some concepts are likely to be ad hoc. This issue will be returned to in ch. 5.

¹⁶ See Rey (1983, 1985, 2009, and 2010) for arguments to this effect.

¹⁷ I do not believe that anything I say in what follows depends on understanding CTM in this specific way. What I say should be compatible, for instance, with a view that takes computational processes to be sensitive to semantic, as well as syntactic, properties. Assuming a more rigid CTM has two advantages, however. One, it makes stating common views in the concepts literature, such as Fodor’s own version of atomism, easier to state and evaluate. Two, it helps to preserve continuity with much of the psychological literature. For issues where it is relevant, the assumption will be eased and noted.

representation of a category of class.”¹⁸ Similarly, Lawrence Barsalou’s description of the human conceptual system captures the prevailing understanding amongst psychologists:

[T]he basic unit of knowledge is the concept . . . Following psychological theories, we assume that a concept, roughly speaking, is knowledge about a particular category . . . Thus knowledge about birds represents the bodies, behaviors and origins of the respective entities.¹⁹

As those theories have had substantial explanatory success over the decades, I take it to be desirable to maintain continuity with those theories and their assumptions wherever possible. That is not to say that one ought to defer to the psychology on all matters regarding concepts and other cognitive phenomena. Indeed, the view I offer later in chapter 5 is revisionary in many ways. Such revisions need to be well-motivated however, and I take mere conflict with philosophical common sense and priorities, or conflict with pre-theoretic intuition, to be insufficient grounds for the denial of well-supported scientific theories.

Second, there has been increasing success in developing theories of the neural realization of concepts.²⁰ Given the plausible assumption that human minds are nothing over and above human brains, and thus that human mental states are nothing over and above human neural states, if there is evidence that concepts are realized in the brain, then there is equal evidence that they are in the mind as well. While capacities may depend on neural states, and while neural states may be related to mere abstracta in a variety of ways, there is little

¹⁸ Komatsu (1992).

¹⁹ Barsalou et al. (2003).

²⁰ See Ashby & Rosedahl (2017) for a neural model for exemplars, Vogele et al. (2001) for theories (in particular theories of the self), Michel (2022) for hybrid theories, and Bowman et al. (2020) for a pluralism including both exemplars and prototypes.

plausibility to the idea that either capacities or mere abstracta can be *identified* with neural states. Likewise, with the other ontological categories that concepts have sometimes been classed under.

Lastly, making this assumption will greatly simplify the discussion that follows, allowing us to bracket complex and contentious metaphysical issues where they are not clearly relevant to the project engaged in. If concepts do turn out to be mental representations and sometime like the Computational Theory of Mind turns out to be right, this I believe greatly strengthens the plausibility of much of what I will say here. However, I do not believe that much I say in the sections that follow depends on the truth of something like CTM, nor on the claim that concepts are mental particulars. As such, reference to mental particulars in what follows could be systematically replaced with talk of abilities, mere abstracta, or whatever entities one wishes, at least if sufficient massaging of the theory is performed.²¹

Even assuming that concepts are mental particulars of some sort does little to answer the question of what exactly they are. There are a whole variety of mental particulars, differing greatly in terms of their functional roles, representational format, semantic content, and so on. Further, as was seen earlier, concepts seem to be expected to play a wide variety of roles, ones which often seem to be in competition and conflict with one another. Luckily, a focus on the theoretical roles of concepts in cognitive science serves to provide the beginning of an answer to this question as well.

²¹ For instance, worries about publicity stemming from the tying together of concept possession and the possession of specific inferential or recognitional capacities, a major theme in what will follow, do not change in substantial ways if one is concerned with the possession of an ability or the possession of a mental particular.

1.3 – Two Oft-Competing Roles

As concepts are theoretical posits meant to explain various phenomena, what it is that they are for is best discerned by looking to the theoretical roles that concepts are expected to play. What exactly are the phenomena concepts are expected to explain, and what exactly are these theoretical roles? While there are many that could be pointed to, there are two main phenomena discussed in the literature that are especially central and worth separating off from the rest:

1. Core cognitive capacities²², i.e. a collection of cognitive abilities of central importance and of which most are found in nearly all thinking things.
2. The capacity to think about aspects of one's environment, i.e. the capacity to represent the world and its parts in thought.

Both of these roles are referenced repeatedly in the work of prominent philosophers and psychologists. For instance, the psychologists Edward Smith and Douglas Medin in their 1981 overview of the psychological literature at the time define concepts as the “critical components of beliefs, preferences, and other cognitive states”, and hold that they are “what provide a cognitive explanation of complex thought and behavior”.²³ A later paper by Medin, Karen Solomon, and Elizabeth Lynch understands concepts similarly:

²² The naming here is inspired by Machery's (2009) talk of core cognitive competencies. I talk of capacities rather than competencies to avoid some of the unwanted connotations of “competency,” in particular how it conjures to mind the ability to reliably execute some task or activity in accordance with a certain standard. My own view allows for concepts which render their possessors more or less incompetent at many of the relevant tasks, as I will ultimately argue that the connection between practical success and concept possession is far looser than has been typically thought. In this I largely agree with Fodor (2004a).

²³ Smith, Medin (1981).

Concepts are the building blocks of thought ... Concepts serve multiple functions, and ... these functions are not independent of one another; rather, they interact with and influence each other ... A concept can be very difficult to define. However ... we will refer to a concept as a mental representation that is used to meet a variety of cognitive functions.²⁴

Another statement of this position is found in Lloyd Komatsu's 1992 review of the development of psychological theories of concepts up to that point:

... a concept is assumed to be the mental representation of a category or class. The contents of such a mental representation ... in concert with certain assumptions about how those contents are processed, have been taken to explain a wide variety of phenomena, including people's knowledge of linguistic relations (e.g., synonymy, antonymy, hyponymy), how people recognize the objects, events, and so on properly labeled by the word (i.e., the extension of the word), how people understand novel combinations of the word with other words, and the inferences people are able to make about an object, event, and so on, properly labeled by the word.²⁵

While both of the two roles are referenced often by both psychologists and philosophers, that is not to say that the two roles receive equal attention or are treated as equally important. Psychologists have largely focused their attention on (1), with some going so far as to deny that (2) is a proper role for concepts at all, arguing that notions like representation and reference play little to no role in scientific theories of concepts.²⁶ Similarly, some philosophers deny that

²⁴ Solomon, Medin, & Lynch (1999).

²⁵ Komatsu (1992)

²⁶ See Hampton (1999).

concepts should have anything to do with the explanation of core cognitive capacities as psychologists understand them, supporting instead general capacities to think about the world or providing a normative ideal for how to think about what the concept is of.²⁷ These positions seem, however, to be in the minority, and most theorists appear to hold that concepts play both roles to at least some extent. That being said, nearly all psychologists and a substantial number of philosophers focus heavily on (1), or at least prioritize it over (2).²⁸ According to these theorists we posit concepts, whatever they turn out to be, primarily with the aim of explaining how their possessors are able to do various things. Core cognitive capacities warranting explanation include categorization, reasoning (both generally and its various more specific flavors, e.g. causal, practical, moral, etc.), and linguistic understanding, among many others. On this way of thinking, concepts are first and foremost cognitive tools, developed so that their possessors are able to accomplish various tasks of practical significance.

Consider a simple case that illustrates how psychologists and many philosophers understand the connection between concepts and core cognitive capacities. A cognitive psychologist may take as their starting point a particular capacity manifested in a subject's behavior, such as a simple capacity to sort objects according to color or shape.²⁹ Such a

²⁷ See Rey (1983, 1985) and Fodor (2004a) and (2008), especially § 2. Rey, in responding to Smith and Medin (1981) and later Smith, Medin, & Rips (1984), holds that what they are offering is not a theory of concepts, but rather a theory of conceptions, i.e. of the representations by which we access concepts.

²⁸ See Prinz & Clark (2004), Peacocke (1992), Machery (2009, 2010), and Millikan (2000) for arguments that a central focus for theories of concepts ought to be explaining some set of practical capacities had by concept possessors. See Fodor (1995, 2004) for arguments against this approach.

²⁹ I will take categorization, recognition, reidentification, and sorting capacities to be equivalent in what follows. Strictly speaking, the capacity to sort requires only a capacity to discriminate, as opposed to a capacity to categorize, and one may be able to reidentify without being able to categorize much at all. In rare instances where it makes a difference, I will make

capacity, it is supposed, depends upon the subject possessing relevant concepts such as RED, BLUE, SQUARE, TRIANGLE, and so forth, which they deploy while engaged in the sorting task. Psychological explanations of this kind often depend on the relevant concepts having semantic content. For instance, the explanation of the subject's capacity to sort square objects from objects of other shapes will appeal to facts such as that their concept SQUARE is about squares and that it includes information concerning the properties of squares and the relations those properties stand in. Plausibly, in the case of SQUARE, this is the definition that squares are four-sided closed figures with sides of equal length.

This information is stored in or encoded by the concept in some manner.³⁰ This information may be encoded by a complex representational structure consisting of further concepts like SIDE, LENGTH, EQUAL, and so on, functionally (or causally, inferentially, mereologically, etc.) related to one another in the right way. Such information is drawn upon to help guide the subject's behavior in the performance of their task, such as via the subject checking a perceived object's represented properties against their mentally represented definition, often understood as a set of necessary and sufficient conditions.³¹ So, one may

clear what capacity is under consideration and what further capacities are required for its possession.

³⁰ By "encoded" in this case I mean simply that information in the world which is available to the organism is converted into a format which can be made use of by the mind. And by "information," I mean information in a loose, rather than a technical, sense. So, for instance, the information carried by smoke regarding fire may be taken in and converted via perceptual channels and then converted again into a conceptual format, where the information is stored in the structural/functional relations found within/between concepts, e.g. between FIRE, SMOKE, and CAUSE. Compare to how the pits and lands of a CD or the pattern of magnetization on VHS tape encodes digital audio and an analog audiovisual signal respectively. I use "encode" rather than more traditional words like "contain" to avoid some of the connotations associated with the container/content model.

³¹ Another possibility is so-called "neo-classical" definitions which encode only necessary properties of the things that fall under them. See Jackendoff (1983), pg. 121.

categorize some perceived shape as a square by recognizing that it has properties like being four-sided which are specified by the concept. This connection between the explanation of core cognitive capacities and conceptual structure will play a major role later on, particularly in chapter 4, as it forms the basis of many arguments for molecularist theories of concepts and poses major issues for their competitors, such as a conceptual atomism.

Psychologists focus so much attention on explaining core cognitive capacities that it is rare to see any engagement at all with the issues concerning and surrounding (2). Psychologists tend to simply assume that concepts have the content of the words which express them, e.g., WATER has the content that it does because its possessors express it via the word “water.” For instance, the psychologist James Hampton says:

How does the psychologist provide an account of what these "concepts" are concepts of? How is the link to the external world established? Of course this question has rarely been considered by many researchers, (which is why philosophers have been at pains to point out the difficulty). The experimenter is already typically a speaker of the same language, and so assumes that the participant in the experiment understands the term in the same way.³²

The psychologist Barbara Malt also discusses this assumption, giving much the same answer as Hampton:

The prevailing assumption has been that many important concepts can be easily identified because they are revealed by words – in fact, for many researchers, the words of English. English nouns such as *hat*, *fish*, *triangle*, *table*, and *robin* are used to

³² Hampton (1999).

identify concepts in work encompassing not only the adult concepts literature but developmental work (e.g., Carey, 2009), computational models (e.g., Rogers & McClelland, 2004), conceptual combination (e.g., Hampton, 1997), and neuroscience (e.g., Mahon & Caramazza, 2007).³³

Conversely, when the experimental subject is asked how they think about water (and thus when they hear the psychologist's use of "water") the assumption is that the concept retrieved must be a concept of water. In a similar vein, they often assume that in experimental tasks involving perception (such as categorizing objects according to visually available properties), the concepts retrieved during these tasks correspond unproblematically to the stimuli used. For instance, if the subject sees a picture of an apple and hits the button indicating a perceived category match, then the assumption is that it must have been APPLE which was retrieved and used to guide that categorization.³⁴ When psychologists deviate from these assumptions, the common view endorsed is that the content of a complex concept is held to be determined in an internalist manner from the content of its primitive constituents and their manner of combination, e.g., BACHELOR refers to bachelors because only bachelors satisfy BACHELOR's constituents UNMARRIED, ADULT, and MALE.

³³ Malt et al. (2011).

³⁴ This methodology is not only acknowledged but explicitly endorsed in Hampton (1999): "The thinker's concepts are individuated using behavioral criteria, such as whether they will judge that it is appropriate to apply a concept term to describe an object. In the case of non-verbal or pre-verbal thinkers like animals or small children, many other possible ways in which the thinker's behavior picks out a certain type of object, event or situation have been used for the exploration of conceptual understanding. Examples are sequential touching of objects in the same class, release from habituation on moving from one class to another, and the ability to learn classifications through instrumental conditioning and discrimination learning." See Hampton (2020) as well.

By contrast, philosophers working within broadly naturalistic and externalistic approaches to the mind tend to prioritize (2), downplaying (or even discounting entirely) the importance of (1).³⁵ For them, the focus is squarely on the fact that concepts are the constituents of the representations which support our having of propositional attitudes. Any theory of concepts should thus above all else explain how it is that an individual can bear such attitudes towards aspects of their environment. Those who prioritize (2) then place semantics in the foreground and place providing psychological explanations of cognitive capacities in the background. The motivations and expectations of psychologists and those of many philosophers theorizing about concepts then are often starkly different.

That is not to say that theorists who primarily focus on the propositional attitudes, and on concepts' roles as constituents of PA-supporting representations, are not concerned with providing explanations of behavior.³⁶ The capacity to have propositional attitudes about aspects of one's environment is, these theorists point out, just as important (if not more so) for explaining behavior as structured concepts are for explaining capacities such as categorization. When a four-year-old categorizes a cat painted with a white stripe down its back as a skunk, a plausible explanation for that behavior is that the child believes that the animal presented looks like a skunk and believes (implicitly) that looking like a skunk is all that there is to being a skunk.³⁷ When the child a year later refuses to make the same categorization, that refusal can

³⁵ For representative examples, see Fodor (2004a) and (2008), § 2.2; Millikan (2017); and Camp (2015).

³⁶ By "PA-supporting representation" I mean a representation such that it is in virtue of possessing such a representation that one has a given propositional attitude or set of attitudes.

³⁷ By "believes implicitly" I mean nothing more than that the child has the belief, but that the belief is not introspectively available and thus not a candidate for verbal report. Such beliefs may despite this still manifest in behavior, such as in the case of implicit biases whose possessors are often ignorant of their existence.

be explained in a similar manner, by appealing to a shift in the child's beliefs. They no longer believe that looking like a skunk is all there is to being a skunk. They now believe that some deep, hidden property determines category membership instead, a property missing in the case of the unfortunate cat.³⁸ Whether this explanation proceeds through an appeal to the child's beliefs or through something like a concept of the sort posited in psychology, such as a child's folk theory of the biological domain, seems to not be of much consequence.³⁹

Likewise, trusty GreyCat's recognition of the container in which its treats are stored is explained just as well by attributing to GreyCat various beliefs (such as where its treats are located and the appearance of the container) as it is by attributing to GreyCat a prototype or exemplars of the treat container in question. It is true that a focus on propositional attitudes over prototypes and the like expands, rather than contract, our explanatory resources by making available desires and other attitudes whose supporting representations are not as well understood as those for belief. Once one helps oneself to various feline propositional attitudes, one can give an explanation, not just of GreyCat's categorization of the treat container as such, but also GreyCat's meowing and desperate lunges for that container. This seems, however, more a reflection of the comparative lack of understanding of representations of these latter sorts.

There is, I think, something odd in the fortuitous coordination of these two sorts of explanation. The concepts posited by psychologists are meant to explain various kinds of behavior, central among them categorization. However, as we saw it seems that those same

³⁸ Keil (1989). See ch. 9 in particular.

³⁹ This alignment in explanatory role between propositional attitudes and concepts as understood by those in psychology is, I think, telling of a mismatch in how philosophers and psychologists think about concepts. Much of ch. 5 will be concerned with this mismatch.

behaviors are amenable to standard folk and philosophical explanations in terms of the propositional attitudes as well. One can say that GreyCat utilizes a prototype and a similarity matching process to recognize where its treats are stored. However, one can equally well say that GreyCat simply knows what its treat container looks like and can check whether what it is currently looking at appears similar. One immediately pressing question then is what exactly the relation between these two explanatory paradigms is. In particular, there is the question of whether these two sorts of explanations are in competition with one another, whether they are complementary, or whether they are simply explanations of the same sort but each at differing levels of abstraction and complexity.

Often, these sorts of explanations have been taken to be competing. On this line of thought, philosophers and psychologists are positing different kinds of mental representations and offering different kinds of explanations of the same phenomenon.⁴⁰ A psychologist may explain categorization and related behavior in terms of prototypes, exemplars, featural representations, a similarity-matching process, accumulators, and categorization thresholds. Meanwhile the philosopher and the person on the street instead explain the same behavior in terms of having beliefs, how things appear to someone, the noticing of similarities, and the making of judgments. Going a step further, these explanations have also sometimes been taken to be concerning different phenomena entirely. For instance, Machery (2009, 2010) argues that psychologist's concepts are used only for snap judgments, i.e. those judgments made under

⁴⁰ See Fodor (1998), in particular ch. 5 on prototypes and whether they can do the work Fodor expects of the conceptual atoms he favors. See Camp (2015) and her distinction between a philosophical language of thought and what she calls "characterizations" in psychology, each of which can perform many of the same functions.

tight time constraints and/or without much conscious consideration.⁴¹ Background beliefs, those propositional attitudes often factoring in the explanations of the philosopher and layperson, are relied upon only when we make more careful, considered, or reasoned judgments:⁴²

The knowledge that is stored in a concept of x is preferentially available when we think, reason, and so on, about x. So to speak, it spontaneously comes to mind. By contrast, the knowledge about x that is not stored in a concept of x is less available—it does not spontaneously come to mind. The knowledge that is not stored in a concept of x is used only when the knowledge that is stored in the concept is insufficient or inadequate for the task at hand. In such cases, people access their long-term memory in order to retrieve some additional knowledge about x that helps them deal with the task at hand (i.e. some knowledge that is not stored in the concept of x).⁴³

On this way of thinking, there is a clear division of cognitive labor: Concepts are for quick or preliminary categorization and reasoning, in cases where error is either unlikely or of little consequence to the thinker. By contrast, propositional attitudes are for cases where

⁴¹ This is a consequence of Machery requiring that concepts consist only of that knowledge which is “spontaneously retrieved” upon a thinker exercising one of their core cognitive capacities.

⁴² This view presumably results in not just a difference between what is being focused upon (psychologists studying concepts only care about snap judgments; philosophers and laypeople when mentioning concepts only care about considered judgments), but a difference in how the etiology of common behaviors is understood (at least in cases where both sides appeal to concepts in some manner). For instance, GreyCat makes a snap judgment when the psychologist invoke GreyCat’s concepts, but GreyCat makes a considered judgment when philosophers and laypeople invoke propositional attitudes containing the relevant concepts.

⁴³ Machery (2009), pg. 11-12. In ch. 3 and 5 I will discuss whether such a distinction should be drawn.

accuracy is prized over rapidity.⁴⁴ These processes can also be seen as separate but complementary, with concepts providing a kind of first pass that can be overridden by the more definitive processes involving propositional attitudes.⁴⁵

While initially plausible, this account does not fit equally well with all the major theories of concepts. It is arguably most apt when discussing simple concepts, such as prototypes and exemplars. These psychologists do intend to set up their experiments in such a way that the subject's draw only upon the information in the prototypes and exemplars, and not further upon information supposed to be stored in background knowledge. They do this by having the subjects perform their tasks under strict time pressure, or having the tasks concern artificial categories about which the subjects have no existing background knowledge. However, this sort of methodology is not utilized by those working in the theory-theory tradition. These subjects are not under such constraints and their judgments are more considered and thought out. Similarly, it seems unlikely that classical concepts, i.e. definitional concepts, will be ones used in snap as opposed to considered judgments. This is especially so when the definitions become more complex and multi-part, where quick checking of one's target against the definition is unlikely.

⁴⁴ This way of thinking is evocative of so-called system 1 and system 2 approaches to cognition, where system 1 processes are for unconscious, rapid reasoning and system 2 processes are for conscious, slower reasoning. See Frankish (2010) for an introduction to this and related approaches.

⁴⁵ This would not necessarily be the same as what is claimed by hybrid approaches. On those views, concepts themselves are divided into distinct parts, some used for quick reasoning and others for more considered reasoning, e.g. a concept may consist of a prototype for rapid but inaccurate categorization and a definition for slower, more deliberate categorization, with the definition taking priority in cases of conflict between the two.

An alternative and natural suggestion is that the explanations that psychologists give in terms of concepts are explanations of the same sort as those given by philosophers and ordinary everyday people. They explain the same phenomena in roughly analogous ways, just with the psychologists offering explanations with a greater degree of specificity and detail. On this way of thinking, one plausible account of what psychologists are doing when they theorize about concepts is that they are investigating and offering theories of the representations which underly and support our having of propositional attitudes. This is rather than them theorizing about a sui generis form of mental representation, as they often taken to be. This is also rather than them theorizing about anything that looks like the constituents of thought, i.e. without them theorizing about what philosophers are most inclined to call concepts. Philosophers and the folk are, in contrast to the psychologists, offering explanations at a much higher level of abstraction. This suggestion I will explore later, when in chapter 5 I distinguish between conceptual atoms and conceptual structures, with conceptual structures sharing much in common with what are commonly called beliefs while conceptual atoms are more akin to the concepts of philosophers.⁴⁶

Up to now I have been writing as if there is a fairly deep, perhaps unbridgeable, rift between how philosophers and psychologists approach the study of concepts. However, this bifurcation is, admittedly, a tad simplistic, even perhaps reductive. There are many philosophers (see Peacocke, Prinz, Machery, and Millikan for representative examples) who

⁴⁶ While this will be explored in more detail later on, it is worth saying at the outset that what I will term conceptual structures may be closer to what philosophers sometimes call “conceptions, rather than beliefs, where conceptions are something like structured complexes of closely related beliefs (such as beliefs about a common subject matter or beliefs that tend to be utilized in the same task).

tie concepts and recognitional/inferential capacities tightly together⁴⁷ Likewise, there are some psychologists (unfortunately, rather few it seems) who take seriously the question of how concepts can acquire semantic properties and the requirements for their supporting propositional attitudes.⁴⁸ Even so, it is not uncommon to see the literature divided in this way (even if only to then argue the division is misguided), and for our purposes here it is useful to paper over much of the complexity in these debates.⁴⁹

What is important to keep in mind is that there are these two broad approaches to understanding concepts: (1) as constituents of PA-supporting mental representations, and (2) as the mental representations which support core cognitive capacities. Theories within each approach may vary from one another in a variety of ways. And many theorists utilize both approaches simultaneously, with these theorists blending and integrating them in a variety of ways. Nevertheless, there has been a tendency for the form that these theories take to be heavily influenced by which approach is followed most closely and by which issue receives the greatest focus. It is not an accident that the shape of a theory of concepts is determined in large part by which of the two explanatory roles, supporting PAs or supporting core cognitive capacities, is seen as the primary. if not the sole, explanandum for a theory of concepts. The effects of this difference in focus on such theories will be seen continually in what follows.

⁴⁷ In Millikan's case, very tightly. Millikan (2000) defends the view that concepts are abilities to reidentify substances across substantial variation in context and intervening media. See ch. 4 in particular.

⁴⁸ See Carey (2009) for one such example. See ch. 1 and 13 in particular.

⁴⁹ See Camp (2014); Machery (2009, 2010); Peacocke (1992); Millikan (2000) pg. 47; Rey (1983, 1984); and Fodor (2004a) for representative examples of such divisions being drawn.

1.4 – Concept Individuation

While disputes over the metaphysical status and theoretical roles of concepts are important ones to settle, they have not been among the most prominent in recent years. The former has been largely in the background, with a particular answer often assumed at the outset of the discussion. The latter has seemingly in recent years begun to reach something of a consensus. One issue that has garnered more attention lately is one that, at first, seems so abstract as to be of little consequence. This issue concerns the question of what individuates concepts, i.e. what is it in virtue of that a given concept kind is distinct from all other kinds?⁵⁰

As it is stated, answering this question appears to be a simple matter: simply figure out what concept kinds are out there in the world (perhaps lodged in each of our heads) and then state identity conditions for each. This task is made even easier if one accepts a few common assumptions. We may assume, for instance, that we can figure out what concepts kinds there are by looking to the objects of those concepts, here assuming there are as many concept kinds, and perhaps no more, as there are things to be represented in thought. Alternatively, we may look towards language, assuming there are concepts corresponding to each semantically distinct word in natural language.⁵¹ These assumptions, if true, would greatly simplify the project of developing an inventory of the concept kinds that exist.⁵²

⁵⁰ Throughout I will treat talk of kinds and talk of types as more or less equivalent. While I will primarily use “kind” when talking about concepts (as doing so will simplify much of the discussion to come), I will switch to using “type” when intricacies having to do with type and token-identity are impossible to talk around or ignore.

⁵¹ Depending on one’s theory of language, this may result in the same or a greater number of concept kinds than the previous approach, with coreferential words expressing distinct kinds of concepts.

⁵² See Millikan (2000, 2017) for an approach that starts with developing a theory of what kinds exist in the world and then works backwards to what concept kinds there must be. See Malt et al. (2011) for reasons why this second assumption may be false.

Unfortunately, it is far from clear that assumptions such as these are well-motivated. Nor is it obvious that determining how concepts are individuated will prove quite so simple. Upon closer inspection, it is even far from clear what the question is asking, what task it poses for us, and how to go about answering it. It is not even clear, I will argue, what it is that we as philosophers (and perhaps cognitive scientists along with us) are trying to accomplish in raising this question in the first place. In short, the question, though simple enough to state, balloons in complexity once one makes any attempt to answer it in detail.

Any particular concept, much as with everything else in the world, belongs simultaneously to many different kinds. Much as a particular cat belongs to such disparate kinds as cat, animal, household pet, and perhaps a particular breed, any given concept can belong to such kinds as cat-concept, definitional concept, and (on certain views) some kind of neurological event (such as a particular pattern of neuronal activation). For each of these kinds, we can ask what makes something a member of that particular kind, arriving at answers with very different forms and implications for each. What makes a particular cat a cat, for instance, is very different from what makes it an animal or a household pet. What makes a cat a cat is its membership in a particular reproductively established lineage of organisms (roughly, what portion of the phylogenetic tree that cat's parentage places it in). What makes it a household pet is facts having to do with the social practices surrounding and supporting pet ownership, the particular history of that organism, and how it is bound up in those practices. In the same way, what makes something a concept is likely to be quite different from what makes something an observational concept, a prototype, or a neurological event of some specific kind. Further, as with the toy example of the cat and its various kinds, it is likely that our answers to questions about concept individuation will draw upon facts and phenomena from quite varied

domains. From the outset then we should not anticipate, much less require, a neat and tidy answer to the question (or more accurately questions) originally posed about how to taxonomize the conceptual domain.

Untangling and distinguishing the various ways in which concepts can be individuated is necessary if we are to get clear on what exactly is at stake when we argue about the nature of concepts. One way to clarify and give focus the investigation is to distinguish the various things we may be asking when we pose the question “What individuates a given concept?” For instance, when we pose this question, we may have in mind issues surrounding what is distinctive of conceptual representation, as opposed to other closely related mental representation kinds like percepts, motor control plans, syntactic structures, and the like. We may also be concerned with the separate question of what makes is distinctive of a particular functional subkind for concepts, e.g. what makes it an observational, theoretical, or practical concept kind (among many others). Or we could instead ask the same question for structural subkinds, e.g. what makes it a prototype, exemplar, theory component, or definition. Finally, we may be asking what is distinctive of a conceptual semantic kind, e.g. what makes it a concept which represents cats as opposed to a concept which represents dogs, i.e. what makes it the concept CAT as opposed to DOG.⁵³ As a final wrinkle, even this final question is really two questions in disguise. We may be concerned with what makes a concept represent the particular aspect of the world that it does, e.g. dogs instead of cats, a question of its *referential content*.⁵⁴ Or we may instead be concerned with *how* it represents whatever it is it represents,

⁵³ Throughout I will be following standard conventions and naming concepts with small caps.

⁵⁴ I use “aspect of the world” and “world-aspect” in what follows as a catch-all term for anything in the world which may be represented in thought. Thus, “aspect” as used here refers to kinds, particulars, properties, events, states of affairs, locations, times, and so on. I use this

a question of its *cognitive content*, e.g. representing dogs via visually observable properties like size and shape, as opposed to theoretical properties like genetics and lineage.⁵⁵

In summary, we can distinguish at least five questions having to do with concept individuation that may be relevant to my project here:

- (1) How should concepts be individuated relative to other mental representation kinds?
- (2) How should concepts be individuated functionally?
- (3) How should concepts be individuated structurally?
- (4) How should concepts be individuated in terms of referential content?
- (5) How should concepts be individuated in terms of cognitive content?

Developing an adequate theory of concepts will require first determining (i) which of these questions are relevant to a complete account of concept individuation, (ii) what it is that determines which of these questions are relevant, and (iii) what factors we must consider when deciding the answer to each of the questions that turn out to be relevant.⁵⁶ Figuring out (i)-(iii) will require careful consideration of the theories of concepts proposed up until now, as well as the sort of theoretical work that they have been expected to perform. Much of this dissertation will be concerned with just this task. Some theories, as we will see in the coming pages, deny

term, in part, to avoid the typical but misleading association between concepts and kinds or categories of things. While concepts do represent such things, they represent many other aspects of the world as well, aspects often not paid sufficient attention to. Similar terminology is used occasionally by Fodor, such as in Fodor (2004), pg. 19.

⁵⁵ The legitimacy of some of these distinctions may be questioned. I will, for instance, ultimately argue that what philosophers call “concepts” lack cognitive content entirely. *How* a subject represents an aspect of their environment is ultimately a question concerning what I will term conceptual structures, complex representations composed of conceptual atoms, which play much of the same role as traditional concepts in philosophy.

⁵⁶ A fourth consideration is whether all of the questions above are substantive, or whether some or even many of them are matters of mere theoretical bookkeeping. See Margolis & Laurence (2007) for discussion of the latter possibility, in particular with regards to questions concerning structural kinds for concepts.

the relevance of some of these questions, such as many traditional atomistic theories which deny we should ask questions (3) and (5), focusing our attention instead on (1), (2), and (4). While (1) through (5) all correspond to proposed ways of typing concepts (and all are ways concepts have been typed in both the philosophical and the psychological literature), it should not be assumed at the outset that all these ways of typing will end up being appropriate or useful for the purposes of developing an adequate theory.⁵⁷

Why though should anyone concern themselves with questions of concept individuation? Psychologists seem to get along just fine without considering the question, making a good degree of progress in spite of their lack of concern with the issue. The answer, in short, is that it is not possible to develop a theory of concepts without committing oneself to a particular picture of concept individuation. This has seemingly gone largely unnoticed in much of psychology but was realized early on in the relevant philosophical debates.⁵⁸ The reason, in all likelihood, is that few psychologists seek to reconcile their own theories with the theories of philosophers. If they attempted to do this, they would quickly realize the great difficulties impeding success in this endeavor, difficulties owing to the quite different approaches that philosophers and psychologists take with regards to issues of concept

⁵⁷ Talk of “typing” I take to refer to a specific human activity whereby we develop taxonomies of some domain for theoretical or practical purposes, e.g. the Linnean classification for biological species is a particular way of typing the biological domain. Individuation, by contrast, has more to do with the nature of the kinds that exist in the world. Ideally, the ways in which we type the world will result in taxonomies that correspond to explanatorily fruitful kinds with known conditions for type-identity.

⁵⁸ In psychology, it is rare to see any direct engagement with issues regarding either the individuation of concepts or issues surrounding the determination of their semantic properties. A rare exception is Carey (2009). Most debates in the psychological literature concern instead the structure of concepts, taking many of their semantic properties for granted.

individuation. As a result, the seeming importance of concept individuation to theorizing about concepts does not present itself to psychologists in so stark a form as it does to philosophers.

While one's theory of concepts has been long thought to commit one to a particular account of concept individuation, the reverse has seemed to many philosophers to be true as well: How one answers questions concerning concept individuation seems to determine much of one's other positions on concepts.⁵⁹ This includes everything from what is required to possess a concept, which/how many concepts are innate, what role concepts play in our psychological theories, how/if concepts are structured, whether concept possession can make possible certain kinds of a priori knowledge, and so forth. Taking one prominent example, Fodor has argued that if concepts are structurally simple and individuated in terms of their referents, a view he terms *concept atomism*, then it follows that all (or nearly all) lexical concepts, i.e. concepts expressible by a single word in natural language, are innate.⁶⁰ Conversely, it has been argued that if concepts are structurally simple, then concepts will be unable to explain the very cognitive capacities they are posited to explain.⁶¹

⁵⁹ Fodor (1995) is explicit on this point, stressing the foundational importance of questions concerning individuation, in particular the tight connection between concept individuation and concept possession.

⁶⁰ See Fodor (1975, 1981) for the arguments for radical concept nativism, and see Laurence, Margolis (2002) for possible ways to understand concept acquisition that do not run afoul of Fodor's arguments. Lexical concepts can be contrasted with phrasal concepts, i.e. those concepts expressible by a single phrase. Note that as stated these definitions are inadequate. A concept expressible by a single word in English may only be expressible by a phrase in another language, say Japanese, and vice versa. See Millikan (1997, 2017) for reasons to think that a single type of concept, relative to a single natural language, could have both lexical and phrasal concept tokens depending upon the context and the mode of acquisition, e.g. I may have a concept of RED SULFUR either because I (1) conjoin my concepts RED and SULFUR to form a complex concept RED SULFUR, or (2) have a simple concept RED SULFUR which bears the right content-determining relation to the world.

⁶¹ See Prinz (2002), pg. 99-100, and Laurence & Margolis (1999) arguments to this effect and discussion of their merit.

One goal throughout this dissertation will be to argue that much of this division is overblown. While there are indeed differences between theories as a result of how they understand concept individuation, the differences are not as grand as is often thought. That is not to say that there are no consequences for choosing one account of concept individuation over another. The differences can be far-reaching and have real effects on how both philosophical and psychological investigation ought to proceed. But nothing as profound as the truth or falsity of radical concept nativism or the possibility of a priori knowledge follows from one's position on issues of concept individuation alone. Fodor's arguments for radical concept nativism, for instance, depend not just on his atomistic theory of concept individuation, but on ancillary claims about concept acquisition. In particular, the argument depends upon the view that concepts must be either innate or acquired via a process of hypothesis testing and confirmation. If other methods of concept acquisition exist (and we have good reasons to think they do), then it is perfectly possible for one to accept Fodor's concept atomism while avoiding the less plausible radical concept nativism to which it is often attached.⁶² Likewise, arguments that concepts must be structured if they are to explain cognitive capacities assume both that (1) concepts are in the business of explaining such capacities in the first place, and (2) that structure needs to be individuating for such explanations to be adequate for the purposes of psychological theorizing. Many of the supposed far-reaching implications of theories of conceptual individuation depend upon background assumptions regarding concepts such as these. Such assumptions and the possibility of less explored theoretical alternatives allows for far more wiggle room in theorizing about concepts than is sometimes thought.

⁶² See Margolis (1998).

A second goal in what follows will be to make clear exactly what the differing proposals regarding concept individuation actually claim and entail. This task runs throughout all that follows but will receive particular attention in chapter 4. An unfortunate feature of much of the recent discussion on concept individuation, in particular when it comes to atomist proposals, has been an unintentional though persistent misconstruing and misunderstanding of the positions on offer. To take on example, it is unfortunately common to see atomism dismissed off hand due to its association with radical nativism or due to explanatory concerns like those mentioned above that the view only suffers from if interpreted in a simplistic manner. That is not to say that these worries and dismissals do not reflect genuine problems with the theories in question. Rather, it is that debates, particularly when occurring across disciplinary divides, often involve a pervasive underestimation of the complexity and nuance of the competing theories, as well as the theoretical resources available to each side.

These two broad goals are related, insofar as getting clear on exactly what is proposed by all sides in this debate will serve to reduce the perceived distance between them. By getting clear on what the proponents of these views are claiming and their more immediate consequences, I hope to show that these theorists are engaged in much the same research program as one another.⁶³ They are not, as is sometimes claimed, simply talking past one another or concerned with wholly different phenomena.⁶⁴ There are real disagreements; that cannot be denied. But the gulf between all sides is not so vast and unbridgeable that one can

⁶³ In particular, I hope to assuage the worry that philosophers and psychologists are concerned with wholly different phenomena when they study concepts. See Machery (2009), pg. 45-37, for a particularly clear statement of this worry.

⁶⁴ Machery (2009, 2010).

ignore the work being done by any one group on the grounds that they are simply theorizing about something other than what one is oneself concerned with.

1.5 – Conceptual Structure

While concept individuation is perhaps the most prominent locus of debate in the philosophy of concepts, there is another which comes close to matching it in prominence in the literature more generally. An issue of nearly equal importance and centrality is whether concepts are structured and, if so, what sort of structure they have. In fact, these issues are often thought to be intertwined, as will be seen soon. Further, while debates about the individuation of concepts are dominant in philosophy, debates about the structure of concepts hold a similarly central place in the psychological debates.

It is important to be careful when discussing the notion of conceptual structure, as structure is often more broadly construed in the literature than may be expected. What “structure” often conjures to mind is some individual thing being complex, able to be broken down into parts which stand in purposeful relations to one another. Understanding conceptual structure in this way, concepts are decomposable into parts, the parts themselves usually conceptual (or if not, at least representational). For instance, the concept CAT on this way of thinking would be a complex mental representation decomposable into component concepts such as ANIMAL, FURRY, MEOWING, and so on. These latter concepts would, in turn, be decomposable into even simpler concepts, with those simpler concepts themselves decomposable, and so on. This process would ultimately bottom out in primitive concepts,

which are either perceptually-based or innate.⁶⁵ Structure of this sort, where a concept is complex and articulable into parts can be termed *internal structure*, as it deals with the internal parts of a concept and their relations to one another.

Clear examples of internal structure can be found in classical, i.e. definitional, approaches to concepts in philosophy, as well as more contemporary approaches in psychology, such as the prototype and exemplar theories. Under the classical theory, concepts are structured as sets of necessary and sufficient conditions. Possessors of a given concept thus possess a mentally represented definition which captures all and only those world-aspects which fall within the extension of that representation. Thus, the concept BACHELOR would have as components concepts such as MAN, ADULT, and UNMARRIED, with these components able to be further broken down (UNMARRIED into NOT and MARRIED, for example). On prototype and exemplar approaches, concepts decompose into feature representations (roughly, representations of properties). So, a concept such as NURSE could be broken down into concepts of the typical or cue-valid⁶⁶ properties of nurses, such as FEMALE, WEARING A WHITE OUTFIT,

⁶⁵ By “perceptually-based” I mean the concept’s content is determined by its relations to perceptual representations, rather than its relations to other concepts. I do *not* mean the stronger claim that the concept is in a perceptual format or is “copied” from a percept, as in the empiricist tradition. There are, however, contemporary theory of concepts (Prinz 2002), in particular ch. 5-7, that make this stronger claim, as do theories of concepts (or ideas) in early modern British empiricism. See Barsalou (1999) and Barsalou et al. (2003) for another neo-empiricist proposal.

⁶⁶ Cue-valid properties are those properties which, if possessed, make it highly likely that something falls under a given category. Typical properties, by contrast, are those which it is highly likely that an arbitrary member of a category possesses. For instance, four-leggedness is a typical, but not a cue-valid property of dogs. If something is a dog, it is highly likely that it has four legs. However, if something has four legs, it is not especially likely that it is a dog, as most four-legged things are not dogs. Cue-valid properties are then, unlike typical properties, highly diagnostic of membership in a given category. Cue-valid properties will be covered in more detail in ch. 2.

WORKS IN A HOSPITAL, and so on.⁶⁷ These feature representations, much like on the classical theory, would themselves be decomposable into more basic representations. On such approaches, the representations of properties are literal parts of the representations of the world-aspects which are represented as possessing those properties.

Another sort of structure is that found in the conceptual, inferential, and functional role theories which tend to be more popular in philosophical circles.⁶⁸ On such views it is not (or at least need not be) the case that concepts are complex in the earlier sense of being built up out of simpler concepts and other representations. Rather, on these views concepts essentially stand in causal/inferential relations to other concepts. It is by standing in such relationships that a concept both secures its status as a concept as well as acquires the semantic properties that it possesses. Thus, while a concept like CAT may not have literal parts on this understanding of structure, it essentially bears causal/inferential relations to other concepts such as ANIMAL, FURRY, MEOWING, and so on. It is by virtue of it standing in such relations that, for instance, a subject is disposed to token CAT in the presence of recognized instances of ANIMAL and vice versa. Structure of this sort can be termed *external structure*, as it deals with the relations between a given concept and other representations external to it.⁶⁹

While this approach is found most commonly among philosophers, it is also common to at least one prominent group of psychologists, namely proponents of the theory-theory of

⁶⁷ Phrasal concepts (those most naturally expressible via a phrase in natural language) are often understood as complex concepts of a kind with lexical concepts. The difference between the two is often thought to be solely a matter of the degree of complexity they exhibit, with phrasal concept more complex by virtue of being built up from lexical concepts.

⁶⁸ See Block (1987) and Peacocke (1992).

⁶⁹ The distinction between internal and external structure corresponds to what Laurence and Margolis (1999) call the containment and inferential models of concepts respectively.

concepts. While theory-theorists are notoriously vague about just what concepts are on their view, the rough idea is that concepts get their content in a way similar to how terms in a scientific theory get theirs (or, at least, how theory theorists believe they do), i.e. by occupying a particular place or role within the theory of which they are a part.⁷⁰ Concepts then are the constituents of mentally represented theories, mostly individuated by and having their semantics determined by something like their theoretical role. To what extent this approach differs from the sorts of role theories common to philosophy need not concern us here.⁷¹ For our purposes, I lump theory-theorists in with causal/conceptual/functional role theorists in holding that concepts are individuated, at least in part, by their inferential/causal/functional relations to other concepts, i.e. by how they are externally structured.

What is common to theorists who appeal to either kind of structure is the idea that concepts are individuated, at least in part, by their relations to other concepts and/or mental representations. Those focusing on internal structure theorize about the mereological relations between conceptual representations, while those focusing on external structure theorize about causal/inferential/functional relations.⁷² On this way of thinking, it is essential to a concept that it stand in the relevant relations. Something is not the concept CAT unless it stands in relations to concepts such as ANIMAL, FURRY, and MEOWING, whether these be part/whole relations or causal/inferential/functional relations. Further, the two kinds of structure need not be mutually

⁷⁰ That is not to say this is how scientific terms in fact acquire their semantic properties. Rather, what is important is that this is how theory-theorists understand scientific theories and terms, and that this understanding and the analogy with scientific theories generally has strongly influenced their theorizing about concepts.

⁷¹ See Carey (2009) for reasons to think that theory-theorists are committed to conceptual role semantics. These issues will be discussed in more detail in § 2.3.

⁷² Though those focus on internal structure will likely think that the concepts that stand in mereological relations also stand in causal/functional/inferential relations as well.

exclusive; there is nothing incoherent in a theory that holds that concepts are individuated both by their parts and their relations, as well as their causal/inferential/functional relations to other external concepts. For instance, one could hold that a concept such as TRIANGLE is the concept that it is *both* because (1) it is composed of concepts such as SHAPE, THREE-SIDED, and so on, as well as (2) it stands in inferential relations that disposes its possessor to infer the presence of triangles on the basis of perceptual experiences representing something as being triangular.

That being said, it is often unclear how important this distinction really is and whether the psychological explanations corresponding to each sort of structure are meaningfully different from one another. In fact, it may be that any theory that holds that concepts are internally structured can be reframed as one where concepts are externally structured and vice versa. Theories in the theory-theory tradition can and have been understood as conceptual representations in their own right, with concepts understood as literal theory components. Likewise, some models of prototypes understand them as sets of primitive feature representations bound together in an inferential network.⁷³ Further, it is not clear that the mereological relations highlighted when talking of internal structure are what is really relevant. Rather, their supposed relevance may simply be a factor of their reliably tracking the causal/inferential/functional relations which are of inarguable theoretical importance. So, both approaches, those focusing on internal and external structure, may end up appealing to the same relations when giving their explanations of core cognitive capacities. The difference between the two approaches may be a mere difference in theoretical modeling, rather than

⁷³ Prinz (2002), pg. 81-82, considers three possibilities regarding what concepts are under the theory-theory: that (1) concepts are min-theories, (2) that concepts are constituents of theories, and (3) that concepts are distinct from but affected by theories.

anything more substantive, though it is difficult to say how such issues could be decided with any certainty.⁷⁴

For our purposes what is important is not the differences between broadly psychological theories, which often hold that concepts are individuated by how they are built up from simpler concepts, and broadly philosophical theories, which hold that concepts are individuated by their causal/inferential/functional role. Nor will debates about how exactly concepts are structured, i.e. structured as prototypes versus structured as exemplars, factor much in the discussion to follow.⁷⁵ What matters is what is common between the positions, namely the view that concepts are in part individuated by their relations to other *specific* concepts and mental representations. Focusing on this commonality will allow us to abstract away from differences between the two sorts of theories that are irrelevant to the questions under consideration. As such, I will use “structure” to refer to both a concept’s role in a larger functional system and to the concept’s meaningful parts (if any) and the relations between

⁷⁴ One possibility is that the distinction, while not now of any clear psychological significance, will become of greater relevance as our understanding of the neural realizations of concepts grows more sophisticated. If there are complex concepts and those concepts share token-identical constituents (realized by token identical neural structures/processes), then external structure and a kind of network picture of the conceptual system(s) seems more plausible. If those complex concepts share at most type-identical constituents, then internal structure may be more accurate instead. However, this is (at best) highly speculative, and the distinction may remain without any clear relevance even as our neurological understanding progresses.

⁷⁵ My suspicion, to be endorsed and argued for in ch. 3, is that a broadly pluralist approach to conceptual structure is most plausible. In other words, we have conceptual representations which are structured as definitions, prototypes, exemplars, theories, and so on. I agree with Machery (2009, 2010) and Weiskopf (2009a, 2009b) that attempts to pit the various psychological theories against one another have been largely counterproductive, and that there is no one theory that is capable of accounting for all of the psychological evidence that has been produced since the beginning of the current age of psychological theorizing about concepts some fifty years ago.

those parts. When greater specificity is required, I will refer directly to conceptual roles and external structure, and to concept parts/components and internal structure as appropriate.

Finally, it is important to distinguish between two different levels of abstraction from which a single structured concept can be considered. On the one hand, concepts can be considered from a relatively high level of abstraction where only the barest outlines of their structures are apparent. It is at this level where concepts are distinguished by the coarse kinds of structures that they possess, and thus the coarsely grained sorts of information they encode. This is the level at which one can recognize the difference between, say, a definition and a prototype, or between a concept that represents a thing's essential properties and one that represents its typical properties. On a lower, more concrete level the specifics of this structure come into view. There one can see not just that a concept is a definition, but that it is a definition with certain elements, a definition that represents what it represents in a particular manner.

Structure of the first, more abstract sort I will refer to throughout as the "coarse structure" of a concept while structure of the latter, more concrete sort I will refer to as "fine structure." It is important to note that these are not different ways of being structured but rather different levels of specificity from which a single structure can be seen. It is, in that way, similar to the difference between considering a painting in terms of its composition, e.g. that it is a landscape with trees, a mountain, and stream, and in terms of the individual strokes of paint on the canvas, e.g. that this composition ultimately consists of green, gray, and blue paint strokes arrayed on the canvas in a particular way. In both the case of a painting and in that of a concept, there cannot be one without the other, though focusing attention on one often comes at the expense of the salience and one's awareness of the other.

2 – Theories of Concepts

Most psychologists and many philosophers agree that concepts are mental representations which factor in the processes underlying core cognitive capacities such as categorization, linguistic comprehension, and reasoning of various sorts. This general, though partial, agreement about the ontology of concepts and their functional role has not, however, translated into anything approaching a consensus regarding the structure of concepts and the nature of the processes they factor in.

In this chapter I will survey the major theories of concepts, with a focus predominately on the theories found in cognitive and developmental psychology.⁷⁶ These include most prominently the classical, prototype, exemplar, and theory-theory of concepts. The focus will not be on whether the theories are correct, or which theory out of those presented should be adopted. As I will argue in the next chapter, none of these theories in isolation is explanatorily adequate, and a pluralist position of some sort is to be preferred. Rather, the focus will be on getting an understanding of how each of these kinds of concepts is understood in the literature and some of the reasons that motivated their development. Through this, I hope to also further understanding of how psychologists understand the theoretical role for concepts generally.

Throughout the survey of the leading theories of concepts to follow, it should be borne in mind that all theories of concepts consist of at least two central components. One component is what one would expect given our discussion up to this point: a theory of the structure of concepts. Just as important, however, is a theory of the processes that exploit and make use of

⁷⁶ Overviews of the major psychological theories can be found in Komatsu (1992), Laurence & Margolis (1999), Machery (2009), Murphy (2002), and Prinz (2002). A more thorough and detailed overview of the classical, prototype, and exemplar theories can be found in Smith and Medin (1981).

that structure during a thinker's exercise of their core cognitive capacities. A theory of the processes is essential because a theory of the structure of concepts by itself generates no predictions regarding how concept possessors will behave.⁷⁷ Without knowing both what information is encoded by concepts and, in addition, how that information is made use of, one cannot say anything about what categorization judgments, inferences, and so on, possessors of those concepts will be disposed to make.

Lastly, three important caveats must be noted. The first is that while talk of “the prototype theory” or “the theory-theory” of concepts is widespread in both psychology and philosophy (talk which I will, myself, freely adopt), it is better to think of these as broad families of theories or as general approaches to the study of concepts. This is because there is a great deal of disagreement within each approach over even the most basic of assumptions. For example, there is widespread disagreement among prototype theorists with regards to the exact structure of prototypes, the processes making use of them, what kinds of information prototypes encode, and even on occasion the ontological status of prototypes.⁷⁸ What unites prototype theories ends up being a fairly thin common commitment: that at least some concepts, whatever they are, represent some statistically relevant properties of a kind in some general manner and that they factor in processes that involve similarity matching at some stage. Much the same is true for each of the other theories.

The second caveat is that the descriptions of the theories that follow are greatly simplified. When discussing prototype theories I will, for instance, use early versions of the

⁷⁷ Smith & Medin (1981); Armstrong, Gleitman, & Gleitman (1983).

⁷⁸ Rosch, for instance, has come to (or perhaps always did) doubt that prototypes are mental representations and further doubts that psychological theories of prototypes are committed to any particular view on the ontological status of prototypes. See Rosch (1978).

theory where prototypes consist of little more than representations of specific, typical instances of a kind (e.g. a pictorial representation of a pigeon as a prototype for birds) or representations of some of the typical properties of a kind and their respective weights (e.g. a representation that lists properties like winged, beaked, etc. and how typical each is of pigeons). More plausible prototype theories exist that were developed more recently, such as the schemata variant of the view which imbues prototypes with more structure and more complicated processes to exploit that added structure.⁷⁹ Or theories of prototypes that make distinctions between feature and attribute representations, with each feature falling under a broad attribute that is itself assigned a weight (e.g. the feature WINGED under the attribute BODY PART and the feature MAKES NESTS under the attribute BEHAVIORS).⁸⁰ However, I think that the arguments that I present later in this do not turn on such fine-grained considerations. As such, going into that level of detail is, I think, at best an unnecessary distraction and complication. At worst, it risks obfuscating the actually relevant issues at hand.

Lastly, my coverage of the empirical evidence for each of the theories is necessarily incomplete. This includes even major, much-discussed issues that were central to these theories' popularity, such as the prototype theory's ability to explain the importance basic level

⁷⁹ See Cohen & Murphy (1984); Smith & Osherson (1984). Schemata separate their feature representations into dimensions, with each dimension receiving a value and limitations on which values can occupy them. For instance, a schemata for humans may have as one of its dimensions *eyes*, which allow for features like BROWN but not features like SALTY or FURRY. And the dimensions may place restriction on one another's values, e.g. a relation between dimensions *occupation* and *financial status* may be such that if GRADUATE STUDENT is the value for *occupation*, then RICH may not be the value for *financial status*.

⁸⁰ See Barsalou & Hale (1992). An advantage of a more advanced prototype such as this is that it can make sense of subjects weighting certain broad aspects of categories as more important than others for category membership, rather than subjects being sensitive solely to individual features. For instance, it could be that subjects generally assign more importance to the physical appearance of birds when categorizing, caring less about what behavior they exhibit.

categories.⁸¹ One could easily fill hundreds of pages just describing the various studies that have provided evidence for one theory and against another (and several have, in fact, done so).⁸² My concern here is mostly to give a sense of what each of these theories says about the structure of concepts and conceptual processes, rather than wade into debates about the theories' merits relative to one another.

2.1 – The Classical Theory

Perhaps the earliest theory of what we would today call concepts is the classical theory (sometimes also known as the definitional theory).⁸³ The classical theory of concepts holds that most concept's structures encode definitions, i.e. the necessary and sufficient conditions that must be satisfied for something to be that which is represented by the concept.⁸⁴ For instance, on the classical theory a concept like BACHELOR would include as constituents concepts like UNMARRIED, ADULT, and MAN. The only things which would satisfy the concept BACHELOR are those things that would also satisfy each its constituent concepts as well, i.e. nothing satisfies the concept BACHELOR unless it also satisfies UNMARRIED, ADULT, etc. These

⁸¹ Basic level categories are those categories at the highest level of abstraction where category members still share many salient properties. For instance, *chair* is a basic level category because chairs share many (especially perceptible) properties in common. By contrast, the category *furniture* is more abstract, but its members share comparably few properties. *Rocking chair* may have members with more properties in common than those of *chair*, but it is less abstract. Basic level categories, like typicality effects, are correlated with a wide variety of cognitive phenomena, most of which are well-explained by prototype and exemplar theories, and arguably unexplainable on earlier theories of concepts, such as the classical theory.

⁸² The best example of such an overview of the psychological literature on concepts that I know of is Murphy (2002).

⁸³ Proponents of the classical theory (in something like the form considered here) can, arguably, be found as far back as Locke (1689) in philosophy and Hull (1920) in psychology.

⁸⁴ One common way of putting this is that classical concepts represent the necessary and sufficient conditions for something to fall within their extension. However, it should be noted that this way of putting things privileges concepts of kinds over concepts of properties and particulars, an already widely prevalent bias in the literature on concepts.

constituents would, in turn, themselves be definitely structured, with UNMARRIED being decomposable into simpler concepts like MARRIED and NOT, and likewise for most other constituents of the original concept.

The proponent of the classical theory need not, and likely should not, hold that *all* concepts are definitionally structured. Someone who subscribes to both the classical theory of concepts and an internal understanding of conceptual structure may hold that primitive concepts lack definitional structure (or perhaps, for that matter, any structure at all).⁸⁵ Indeed, they would have to on pain of regress, unless some coherency could be given to the idea of infinitely divisible concepts. These primitive components could be either conceptual and innate or perceptual in nature.⁸⁶ It is less clear that a proponent of an external model of conceptual structure is similarly committed to the existence of primitive concepts. They could, for instance, hold that every concept is part of some group or groups of interdefined concepts. While it is too far afield to consider this particular question in any real detail, it can be noted that primitive concepts are, at the very least, consistent with a view of conceptual structure that is both external and classical in nature.

⁸⁵ The internal understanding of conceptual structure, or what is sometimes called the “containment model,” holds that concepts literally contain their constituents as parts, as opposed to the external or inferential model of conceptual structure that holds that concepts are essentially related causally, functionally, and/or inferentially to some specific group of other concepts.

⁸⁶ The assumption that concepts could be ultimately decomposed into primitive perceptual components was common until around the middle of the last century, with its popularity largely waning alongside parallel declines in the popularity of radical empiricist views in philosophy and psychology. Now it is far more common to see at least some concepts thought to be innate, in particular with regards to domains like biology, psychology, and social relations where there is a plausible case to be made for their evolutionary basis (see Margolis & Lawrence, 2013).

The classical theory remained the consensus position in both philosophy and psychology until around the late 60s. Much of its appeal came from the tidy, intuitive explanations it gave for the phenomena of interest to psychologists and the promises it held for philosophical methodology. On the side of psychology, two phenomena in particular stand out: categorization and concept acquisition.⁸⁷ Categorization on the classical theory is a matter of checking that all the properties represented by the concept are possessed by the thing under consideration. In other words, categorizing something using a classical concept requires checking to see if each of the concept's components is itself applicable. Consider a toy example involving the categorization of someone as a bachelor. If upon encountering someone and wondering whether or not they are a bachelor, I can (so the classical theory claims) simply first check whether BACHELOR's constituents (UNMARRIED, MALE, ADULT, etc.) themselves apply. If they all do, then BACHELOR does as well. There is no need to weight each component concept by salience or typicality, no priority ordering between different levels of represented properties, no need for a process of adjudicating between individually plausible but conflicting possibilities, no need for multi-step categorization processes, and so on. There is the simple consultation of the mental equivalent of a list of individually necessary and jointly sufficient properties, mentally ticking the box next to each one present, and making a positive categorization judgment if, and only if, every "box" ends up with a "tick" beside it.

⁸⁷ A recurring theme in the psychological literature is categorization receiving significantly more attention than the other core cognitive capacities. This outsized focus will be apparent throughout our consideration of the other theories in this chapter. This is not to say, however, that there is no literature on the other capacities. Linguistic comprehension, in particular, has received extensive treatment in the field of cognitive linguistics. See Evans & Green (2006), in particular ch. 5-8, for an overview of some of this work.

Concept acquisition receives a similarly simple explanation.⁸⁸ Concepts on the classical view are simple assemblages of conceptual representations of necessary and sufficient properties, with little more in the way of complex structure binding them together. Acquiring a concept requires little more than (1) first possessing concepts for each of the definitional properties, and (2) combining these together into a complex concept via conjunction in the right way. If I already possess the concepts MARRIED, NOT, MAN, and ADULT, then combining them in the right matter (conjoining, for instance, NOT and MARRIED rather than NOT and ADULT) will result in me possessing the concept BACHELOR.

This is, of course, glossing over a great deal of complexity. The phrase “being combined in the right way” does considerable work here. It is not enough for the classical concept BACHELOR to merely contain as constituents UNMARRIED, ADULT, and MALE. A prototype could equally contain such constituents. What is needed in addition is something in the concept’s structure which disposes possessors of the concept to treat those represented properties as necessary rather than, for instance, merely typical, as they would be if they were parts of a prototype. In other words, the concept must not just represent the relevant properties but in addition represent them as being necessary and jointly sufficient. I will, however, set these additional complexities aside.⁸⁹

⁸⁸ See Bruner, Goodnow, & Austin (1986) for a representative account of how concept acquisition proceeds under the classical theory. See Bruner (1999) for a more brief overview of the same.

⁸⁹ Perhaps somewhat oddly, there is nothing stopping prototypes under the prototype theory from representing properties as necessary and/or sufficient. All that is needed is setting the weights for the relevant properties and/or the categorization threshold correctly. For instance, a property could be understood by the subject as sufficient if its weight was set high enough that recognizing possession of that property alone results in passing the categorization threshold and thus results in making a positive categorization judgment. Likewise, a property

One important detail is that for psychologists working within the classical approach, while concepts may encode definitions there is no requirement or even expectation that these definitions be complete or even correct. Knowledge of the correct definition for a kind is then no condition on the possession of the corresponding concept for psychologists. If a person possesses the concept WATER that does not mean that they thereby know that water is H₂O. Nor does their falsely believing that water is, say, one of the four fundamental elements of which all is constituted, count against them possessing the concept. Aristotle is equally a possessor of the, or at least *a*, concept WATER as Kripke is, though if their concepts are classical they do not encode the same definition for water. This is not an assumption unique to the classical theory; psychologists are equally willing to countenance inaccurate concepts of all types. As such, this assumption should be kept in mind throughout the discussion to follow.

Regarding the theory's philosophical appeal, the primary draw was that it provided the underpinnings for a philosophical methodology centered around conceptual analysis. If concepts encode definitions, representing necessary and sufficient conditions for their own application, then an analysis of one's own concepts may provide a means of learning necessary truths about the world a priori. The thought goes that in reflecting upon real or imagined cases, the inferences I make are guided by the concepts I possess and their cognitive content. The inferences I am disposed to make are, if I am truly a possessor of the concept in question or "grasp" the concept to a sufficient degree, ones I can with some confidence conclude are true. Thus, by considering enough cases and cases with enough variety, I can work backwards from

could be understood as necessary if the categorization threshold and the property's weight are set in such a way that it is impossible to reach that threshold without the recognition of that property. This is just one of many ways that the lines between the various theories of concepts can become blurred, a theme that will continue in what follows.

the inferences I am disposed towards to the content of the concepts guiding those inferences, and then further from the content of the concept to the structure of the world itself. If, for instance, I consider a wide variety of Gettier cases and in each of them I refuse to categorize the relevant beliefs as knowledge, then there must be something in my concept KNOWLEDGE that makes me categorize as such. By looking for similarities in the cases considered I may be able to determine which features of these cases my concepts are sensitive to, such as the role of luck or the disconnect between justification and truth, and in that way suss out how my concept KNOWLEDGE represents knowledge as being. From there I can further infer that, because I do possess or strongly grasp KNOWLEDGE (and thus the way my concept represents knowledge tracks how knowledge in fact is), knowledge has those properties my concept represents it as having.⁹⁰

Now the classical theory on its own does not make possible any such analysis. The theory needs to be supplemented with, by my count, at least two assumptions. The first of these assumptions is the view, widely held at the time, that possessing a concept requires or entails knowing the definition it encodes, at least implicitly. Contained within this assumption of course is the further assumption that one's concepts are accurate, that what they represent as the necessary and sufficient conditions for kind membership are, in fact, necessary and

⁹⁰ See Millikan (2000), ch. 8, for discussion of similar issues and how philosophers often fall importing features of the world into their representations while simultaneously exporting features of their representations back onto the things they represent, with the result being a muddling together of both. Regarding perception, Millikan (pg. 113) says, "The passive picture theory projects properties claimed in or by the visaging [visaging is to perceiving what believing is to knowing] onto the inner vehicle of the visaging ... It also projects properties of the vehicle of the visaging into the visaging's content ... The illusion is thus created both that one directly apprehends aspects of the nature of the vehicle of perception in apprehending the visaged object, and also the reverse, that one can argue from the nature of the vehicle of perception to what must be visaged."

sufficient. The second assumption is that one can, via reflection and/or the consideration of one's inferential dispositions regarding real/imagined cases, come to discern the definition encoded by the concept and distinguish it from information that is not stored in the concept (perhaps that being stored in one's background beliefs or perhaps even another concept of that kind). The philosopher interested in using the classical theory to explicate conceptual analysis as a philosophical method cannot, as the psychologist does, allow that the definitions believed by possessors of the concept are wrong. Nor can they believe that the process of discerning conceptual cognitive content from inferential dispositions is a fallible one.

If any one of these assumptions is false, then traditional philosophical analysis of concepts is, if not wholly misguided, at least far more limited than was typically thought. For instance, if concepts are not definitional and represent instead, say, typical properties of category instances, then one loses the possibility of reading necessary aspects of the world off of our thought a priori. If my concept KNOWLEDGE represents knowledge as *usually* not being the product of luck, then I can hardly infer that knowledge *must* have this feature because of how my concept represents it as being. Likewise, if we are not in an epistemic position to learn of the content of our concepts via reflection on either the concept or our inferential dispositions alone, perhaps being unable to distinguish between information that is properly conceptual and that which constitutes mere background belief, then yet again our efforts at analysis will be frustrated. Finally, if our concepts are inaccurate, then even if they encode definitions concerning the world-aspects we are thinking about, the mere fact they represent such and such properties as being necessary says little about whether or not they actually are.

These assumptions are, I believe, all false. Much of what follows in later chapters will be concerned with exactly why they are false and what some of their pernicious effects on the

study of concepts, especially in philosophy, have been. These problems will be covered in much more detail in all the chapters to follow, particularly 5-7. For now, the promise that the classical theory held for philosophical methodology is what is of most relevance.

The second major supposed advantage of the classical theory in the minds of philosophers at the time was that it provided a workable account of reference determination for concepts. On one way of understanding the classical theory, the one most common during the theory's height, a concept's extension is whatever things satisfy the represented definition. That the referent of the concept is determined by a definition (again, assuming that definition is correct) ensures that the concept has determinate referential content. If concepts represented not definitions but instead, like prototypes, something like typical properties, the extension determined would exclude clear but non-standard category instances as well as include things which, while bearing superficial resemblances to instances of the category, fail to actually belong to it. My prototype of bachelors excludes, for example, the pope, while including individuals who remain married but live a lifestyle typical of a bachelor. So long as internalism about mental content was the dominant position among philosophers, the classical theory offered the only kind of conceptual representations that could plausibly determine their own extensions.⁹¹

⁹¹ The theory-theory could have, in principle, provided the same, given its partial focus on representing the essential properties of world-aspects. However, there are two issues. First, internalism was largely out of fashion by the time the theory-theory emerged in the early 80s. Second, as with the other kinds of conceptual representations posited by psychologists, the folk theories of the theory-theory are not accurate to the world and often highly incomplete. Thus, they would not serve to adequately determine their own extensions (or that of their components) in the way that a classical/definitional concept promised to.

Despite its various strengths and its longstanding appeal, the classical theory has fallen on hard times and today has few supporters.⁹² This is because the classical theory suffers from several widely discussed and seemingly unsolvable issues. These issues can be divided into two broad classes: psychological and philosophical. Unsurprisingly, the psychological issues were largely what led those in psychology to abandon the view, while a different, though equally serious, set of problems led philosophers to seek alternative theories around the same time.

Among the psychological issues, the most influential historically was the discovery of typicality effects in the early 70s and the inability of the classical theory to explain them. Typicality effects arise due to subjects understanding some category members and/or properties as being more typical or representative of the category than others. For instance, a cream-colored loveseat will be judged to be a more typical piece of furniture (at least by most subjects in the US) than a Japanese tatami mat or a hot pink 60s-style egg chair. Similarly, each of the latter will be judged to be more typical than an F-15 ACES II ejection seat repurposed for use in a man cave, or an industrial-size wine barrel turned bed.⁹³ The classical theory does not itself explain why subjects make the typicality judgments they do; pigeons, chickens, ostriches, and penguins all satisfy the definition encoded by BIRD to the exact same degree. However, that is not the real problem. Of greater concern is that these typicality judgments are correlated with a wide variety of measurable phenomena. A small sampling of these include:

⁹² Peacocke is likely the most prominent remaining defender of something like the classical theory.

⁹³ Each of these are real pieces of furniture that can, as of the time of this writing, be purchased.

1. Speed of categorization increases with the typicality of the thing being categorized.⁹⁴
2. Speed of producing some category instance in category production tasks increases with typicality.⁹⁵
3. The likelihood that a category instance will be given as an example increases with its typicality.⁹⁶
4. Categorization error rates increase with atypicality.⁹⁷
5. The rate at which category instances are learned increases with typicality.⁹⁸
6. The likelihood a subject infers the possession of further properties increases with typicality.⁹⁹
7. The speed of linguistic comprehension increases with the typicality of the things referred to by a word/words in a sentence.¹⁰⁰
8. The degree to which subjects waver in their category judgments increases with atypicality.¹⁰¹
9. Failures of transitivity in category judgments.¹⁰²

⁹⁴ Murphy & Brownell (1985); Rips, Shoben, & Smith (1973); Rosch (1973); Smith, Balzano, & Walker (1978).

⁹⁵ Battig & Montague (1969). Category production tasks are those where subjects are provided with a category and asked to list or state members of that category.

⁹⁶ Mervis, Catlin, & Rosch (1976).

⁹⁷ Smith & Medin (1981).

⁹⁸ Mervis & Pani (1980); Posner & Keele (1968); Rosch, Simpson, & Miller (1976).

⁹⁹ Rips (1975)

¹⁰⁰ Garrod & Sanford (1977)

¹⁰¹ Meints, Plunkett, & Harris (1999).

¹⁰² For instance, subjects tend to judge both that clocks are furniture and that Big Ben is a clock, while denying that Big Ben is furniture. See Hampton (1982). Note that “Big Ben” actually names the bell inside the clock tower, rather than the clock itself.

The problem for the classical theory is not that it is inconsistent with any of these experimental findings. Rather, the problem is that it both provides no explanation for them and that it is unclear whether it has the theoretical resources to do so.

While typicality effects were highly influential in the shift of psychologists away from the classical theory towards similarity-based approaches like the prototype and exemplar theories, they are arguably not the most serious empirical issue facing the theory. A perhaps more serious issue is simply that there is little to no actual psychological evidence for the existence of definitionally structured concepts. From the perspective of developing scientific explanations of categorization, inferential, and other forms of behavior, the supposed mentally represented definitions of the classical theory, if they exist, appear to be either entirely idle or left unused by subjects.¹⁰³ According to Fodor:

[T]he number and variety of psychological phenomena that definition theories do not predict, and the reliability with which they do not predict them, is about as impressive as anything known in cognitive science. Definitions don't predict the relative accessibility of concepts or the relative difficulty of concept application tasks; they don't predict the order of acquisition of concepts, or of the words that the concepts are expressed by; they don't predict the demands entertaining a concept make on memory or attention; they don't predict relations of conceptual inter-facilitation in priming or related tasks. In fact, to my knowledge, there is no experimental environment that distinguishes the consequences of definitional connections from those of association or

¹⁰³ Kintsch (1974), pg. 230-233.

empirical centrality; not even in what are supposed to be the clearest cases of definable concepts.¹⁰⁴

Taking one example in more detail, how quickly a subject completes a categorization task appears to bear no relation to how complex the relevant concept likely is if concepts are definitions. To take one example, one would expect on the classical view that the categorization of someone as a man would proceed more quickly than the categorization of that same person as a bachelor. After all, in order to check whether a person satisfies the definition of bachelor one would first have to determine that they satisfy the definition for man. Likewise one would expect that error rates for categorizing with a concept like MURDER would be higher than that for KILL, as one cannot on the classical view categorize a given act as an instance of murder without first categorizing it as an instance of a killing. In each case, the increased complexity of the structure and processing for BACHELOR and MURDER versus MAN and KILL should result in longer and more error prone processes. This is not, however, what psychologists have found.¹⁰⁵

Shifting now to the philosophical issues, the most serious problem is that it places a very high demand on concept possessors. If possessing a particular concept requires possessing a concept that encodes a mental definition for some world-aspect, then concept possessors will need to know, for each of the world-aspects they think about in day-to-day life, the necessary and sufficient conditions for something to be that world-aspect. For most people, if not all, this is a demand that is for all practical purposes impossible to meet.

¹⁰⁴ Fodor (1995).

¹⁰⁵ The case here is from Fodor (2008), pg. 26-27.

Another related issue is that for most concepts there simply do not seem to be any definitions to be had, nor do the prospects for coming up with such definitions seem promising. Despite millennia of attempts, there exist few uncontroversial definitions.¹⁰⁶ Examples are countless, with the many failed attempts at philosophical analysis constituting particularly clear examples. Worse, scientific kinds (which one may expect to be more well-behaved) often seem no more amenable to definition than the kinds philosophers often focus upon. This is now widely accepted with regards to biological kinds such as tiger and other species, where there are multiple ways of defining them depending on one's explanatory ends.¹⁰⁷ Even chemical kinds such as water that seem at first to have a clear definition in terms of its chemical structure are controversial and threatened by various edge cases.¹⁰⁸ Some kinds, such as those for species, may be undefinable in principle, with seemingly no non-arbitrary way to mark off exactly where and when on the phylogenetic tree the species first emerged and with proposed definitions in terms of ancestry seemingly circular. If there are no definitions (or even just not enough of them), then demanding that concept possessors mentally represent such definitions is to demand of them the impossible.

While each of these problems may have been resolvable on their own, the discovery of and/growing awareness of them all within the span of a few short years pushed many to begin

¹⁰⁶ Even the stalwart BACHELOR is not as easily defined as often thought. Intuitively, the definition encoded by BACHELOR would inherit all the vagaries of its component definitions. And as the philosophy of gender should make clear, providing definitions of things such as *man* and *woman* is anything but an easy task. Nor do the prospects for defining *marriage* seem much better, especially given the differences in that institution across place and time. There also remain problem cases, such as that of gay men in civil unions, where people often refrain from categorizing them as bachelors, or in cases like that of green card marriages, where people are often willing to categorize the husband as a bachelor. Though these latter cases may be the result of the use of something like a prototype of bachelor.

¹⁰⁷ Dupré (1993), Kitcher (1984), and Mishler et al. (1987).

¹⁰⁸ See Malt (1994), LaPorte (1996), VandeWall (2007), and Weisberg (2005).

looking for other options.¹⁰⁹ The emergence of competing theories, ones not saddled with the problems discussed above and with many successes to their credit, was more than anything what ended the classical theory as a comprehensive and widely supported theory of concepts.

2.2 – The Prototype and Exemplar Theories

As the classical theory's popularity waned, two new theories began to take its place as part of a newly dominant approach to concepts. Their rapid rise in prominence is largely due to these new theories being explicitly designed to accommodate the very experimental results that the classical theory proved unable to and which accounted for its decline.

These new theories focused, not on the defining properties of the classical theory, but on properties that are typically but contingently had by world-aspects. Both prototypes and exemplars are conceptual representations that in one of many ways represent the typical or cue-valid properties of some aspect of the world.¹¹⁰ With the classical theory, all instances of a kind are expected to have the properties they are represented as having by the concept. Insofar as this is not the case, the classical concept is inadequate. By contrast, instances of the kind represented by the prototype may well fail to have some or even many of the properties represented. Similarly, there may be things that have most or even all of the represented properties, but which fail to belong to the represented kind. Further, this may all be understood

¹⁰⁹ There are some who think the problems faced by the classical theory are not as dire as often thought. See Armstrong, Gleitman, & Gleitman (1983), for instance, for arguments that the classical theory is able to accommodate and explain typicality effects.

¹¹⁰ A property P is a typical property of some kind K iff some arbitrary world-aspect W being a K makes it probable that W has P. A property P is a cue-valid property of some kind K iff some arbitrary world-aspect W having P makes it probable that W is a K. For example, while being four-legged is a typical property of dogs (if something is a dog, it is probable that it is four-legged), four-leggedness is not a cue valid property of dogs (if something is four-legged, it is not probable that it is a dog; in fact, it is quite improbable).

by the possessor of the prototype, i.e. they fully understand that these properties are not necessary nor sufficient for membership in the category.

With regards to their structure, prototypes and exemplars are held on these views to be composed of simpler representations of properties, often called features or featural representations.¹¹¹ These features are often, in turn, each weighted in a way that reflects how salient to kind membership the represented properties are understood to be by the concept's possessor. The processes which exploit the structure of these concepts and their constituent features, at least in the case of categorization, involve checking for similarity between the features and the represented properties of the target of categorization. If the target is represented as being sufficiently similar to how the category is represented by the concept, then a positive categorization judgment is made. If the similarity is insufficient, then either a negative judgment is made, no judgment is made, or the subject wavers, depending on how close the similarity is to the necessary threshold for positive categorization.

Prototypes take a variety of forms depending on the specific permutation of the theory one considers.¹¹² In their simplest form, prototypes are representations of typical category instances.¹¹³ For instance, my BIRD prototype may consist of a representation of a robin or some other common bird and its properties. Another common form prototypes may take is a so-called feature list, a representation of the typical or cue-valid properties of some kind with each

¹¹¹ "Feature" is used in such a way that it can mean either (1) the representation of some property or (2) the represented property itself. I will exclusively use the word in the first sense.

¹¹² Examples of prototype theories can be found in Smith, Shoben, & Rips (1974), Rosch & Mervis (1975), Murphy (1993), and Hampton (2006).

¹¹³ The use of "prototype" is often inconsistent in psychology. Sometimes it refers to the mental representation of a typical category instance, while other times it refers to that instance itself. I will exclusively use the word in the former sense throughout what follows.

feature weighted according to its salience. Here my bird prototype would consist of a list of features like FLIES, HAS WINGS, MAKES NESTS, and so on. More complicated versions of the view may separate features into different categories called attributes, with these themselves weighted by salience. On this view a bird prototype may list HAS WINGS under an attribute like BODY PARTS, weighing these aspects of birds more or less heavily than others such as BEHAVIORS and the features like MAKES NESTS that fall under them.¹¹⁴

Exemplars¹¹⁵ are either (1) representations of particulars, or (2) sets of such representations.¹¹⁶ Understood in the first way, exemplars may seem little different from prototypes, provided the latter are construed as representations of prototypical category members. One difference is that there is no requirement that exemplars represent typical category instances.¹¹⁷ One likely, for instance, possesses an exemplar of their own mother and uses that to categorize and reason about her. However, it may also be the case that one's mother is unrepresentative of mothers more generally (e.g. being atypical in age for one's culture), possessing few if any of the typical or cue-valid properties of the kind.

¹¹⁴ See Barsalou and Hale (1992) for an example of this sort of theory.

¹¹⁵ Similar to the issues with the prototype theory, use of "exemplar" is often ambiguous in the psychological literature. Sometimes "exemplar" refers to the mental representation of a particular category instance, and other times it refers to the particular instance being represented.

¹¹⁶ The exemplar theory originates in the work of Medin & Schaffer (1978), and Brooks (1978). See as well Medin & Schwanenflugel (1981), Hintzman (1986), Nosofsky (1986), and Estes (1994). See Murphy (2002), ch. 4, for a detailed overview of the literature.

¹¹⁷ Another difference is found, not in the structure of the concepts and what they represent, but in the processes that make use of them. Prototype theories tend to posit similarity-matching processes with additive functions, while exemplar theories tend to posit multiplicative functions. This has the result that prototype theories tend to favor categorization based on a high degree of overlap in heavily weighted features without too much emphasis on failure to possess some small number of features. The exemplar theory by contrast is far more likely to predict a negative categorization judgment in cases where a small number of heavily weighted properties are not shared in common.

Given the outsized attention to concepts of kinds over concepts of particulars, it is not surprising that it is exemplar sets rather than the individual exemplars composing them that have been most discussed. Like prototypes, exemplar sets represent the typical or cue valid properties of category instances. The representation of these properties on the exemplar theory however comes, not from the representation of a particular typical instance or via directly representing such properties in a feature list, but rather from the overlap in the properties represented as being had by the many exemplars in the set. A prototype, for example, represents feathered as a typical property of parrots either by representing a typical parrot which happens to be represented as being feathered or via representing featheriness directly as part of a list of typical properties of parrots. An exemplar set, by contrast, represents featheriness as typical of parrots by representing multiple parrots (typical or not), a majority of which are represented as feathered. It is the difference between thinking that parrots are feathered because (1) Polly is a typical parrot and Polly is feathered, and (2) Polly, Rio, Coco, and Sterling are all parrots and Polly, Rio, and Coco are feathered (Sterling, perhaps, suffering from severe molting).

The literature on prototypes and exemplars, in particular how their theories explain typicality effects, is extensive. And their explanations of core cognitive capacities share a good deal in common. In categorization tasks, for instance, both involve representations of typical/cue-valid properties involved in similarity matching between the concept and the target of the categorization task. Explanations of such representations and processes could fill an entire book, and many such books already exist.¹¹⁸ I will focus on just one of many such explanations, that given by the prototype theory for the faster and less error-prone

¹¹⁸ See Smith & Medin (1981) for one especially in depth example.

categorization of typical category instances relative to atypical ones. In doing so I hope to give at least some sense of the shape that these explanations tend to take.

For simplicity's sake, I will focus on a simple feature list version of a prototype. Such a concept could be visualized as followed. Suppose the prototype represents some world-aspect W . F_1 through F_5 are the constituent features of the prototype, each representing some property W is represented as typically having. The values to the right of $F_1 - F_5$ are each feature's corresponding weight, which itself tracks how salient the property represented by the feature is to the subject. Lastly these features are listed in order of their respective weights, with more heavily weighted and thus more salient features appearing first.

F_1	0.7
F_2	0.5
F_3	0.3
F_4	0.2
F_5	0.1

Fig. 2.1

During a categorization task, the subject compares the features found on this list against the categorization target (in this case W) and its properties.¹¹⁹ When there is a match between a feature on the list and a represented property of the target, the weight of that feature is added to an accumulator (suppose A is the sum of the weights of the features the subject takes W as

¹¹⁹ Often this is described as comparing the prototype against yet another representation of the categorization target. It is unclear what this other representation is meant to be. A natural suggestion is that it is perceptual. However, if this is the case, the perceptual representation would have to be capable of representing rather high-level properties, which it is not clear such representations can do. In ch. 5, I will present the outlines of an alternative to this model.

having). This accumulator has a threshold that once passed results in a positive categorization judgment. If not passed, then either a negative judgment is made or no determinate judgment is made, depending on how close A is to passing the threshold. Suppose in this case that the categorization threshold for the prototype is roughly 1.0 (here “~” marks the value that follows as an approximation, rather than standing for negation). A categorization task using a prototype may then be diagrammed as such:

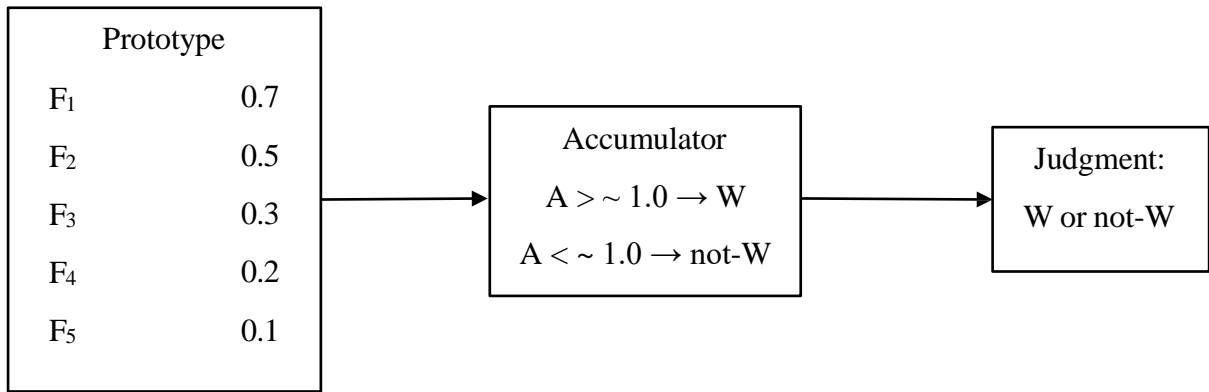


Fig. 2.2

Suppose for the sake of simplicity that the process of comparing the prototype against the target proceeds in a serial manner, i.e. the features are compared one by one in a linear order.¹²⁰ Suppose further that the target is only compared against a single prototype, as opposed to the multiple potentially relevant prototypes that would be simultaneously drawn upon in a

¹²⁰ While some prototype models do conform to this, there is of course no necessity that they do so. The feature comparisons could be carried out in a parallel fashion, which may result in differences in things such as the speed of categorization judgments, e.g. parallel similarity-matching may result in reaching the categorization threshold more quickly, which would result in faster judgments.

realistic case.¹²¹ In this example, the categorization threshold will be passed after just the second of these feature comparisons. When F_1 and F_2 are found to be present, their respective weights are summed to reach 1.2 and the categorization threshold of 1.0 is easily passed. The target is then judged to be W and judged rather quickly because it possesses both F_1 and F_2 , both highly weighted and thus taken to be highly typical properties of W .

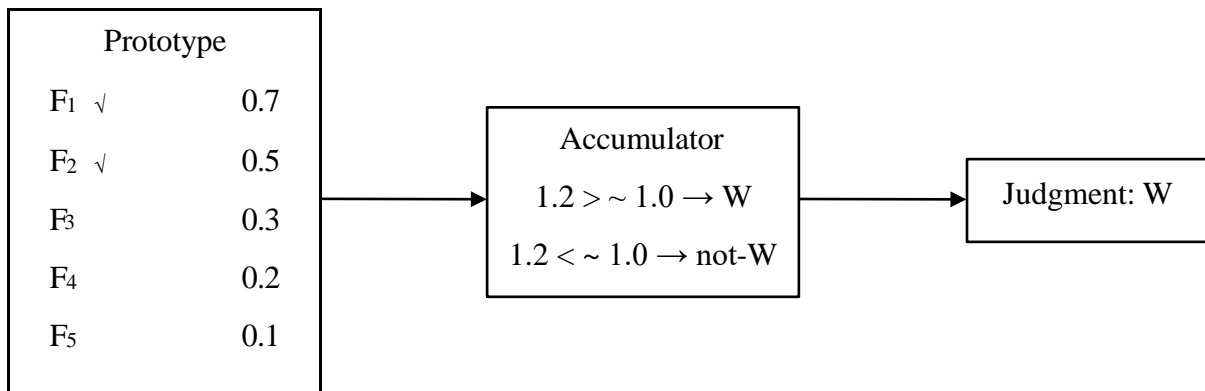


Fig. 2.3

Suppose that the target was something else, an instance less typical of the kind. In this case the target lacks the highly typical and salient feature F_1 and possessing F_2 alone is not sufficient to pass the relevant threshold. Only when F_3 , F_4 , and F_5 have been matched as well and their respective weights added to the accumulator will the threshold be passed and a positive judgment made. Thus, while the more typical category instance of the previous example only takes two feature comparisons before being categorized, this more atypical instance requires a total of five such comparisons.

¹²¹ For example, an animal seen in the distance on a dark and cloudy night may cause the retrieval of prototypes of cats, dogs, raccoons, and other such animals. And it may be that positive judgments result for each, in which case I would have to likely pull from additional cognitive resources to eliminate the resulting inconsistencies.

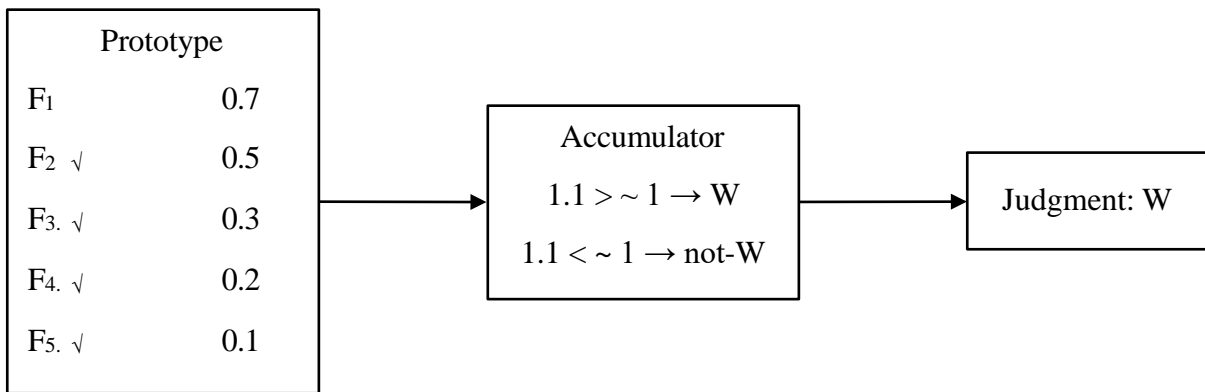


Fig. 2.4

What makes the first target more typical than the second is that it possesses more and more properties taken by the subject to be typical of the category. This is reflected in the fact that the sum of the weights of all its constituent features is greater than the sum of the second's. This increased typicality for the first target results in shorter categorization times because the processing of the relevant prototype involves fewer steps, in this case feature comparisons, than the second. Further, the rate of error is lower because with fewer steps there are fewer opportunities for such errors to arise. While these examples are admittedly simplistic, they are representative of the early models of prototypes and their processes. While they have long been surpassed by more complicated and able models, the outlines remain much the same.

If concepts are understood as psychologists understand them, namely as mental stores of information that are exploited in tasks like categorization and the comprehending of linguistic utterances and inscriptions, then the emergence of similarity-based theories like the prototype and exemplar theories should seem natural or even inevitable. Reflection on ordinary experience reveals that we often do rely on comparing things to typical instances of the relevant kinds or by looking for diagnostic but inessential properties when we categorize things. A task

as simple as recognizing the items on a grocery list while out shopping relies on looking for such typical properties as the size, shape, and color of produce or for the characteristic shapes, logos, and names on food packaging. None of these are properties are essential; a company may change its logos or packaging in a way that frustrates my attempts at easy recognition. They are, however, often the most useful properties for me to base my attempts on, and it should thus be unsurprising that we possess concepts and processes of the sort that prototype and exemplar theories posit.

2.3 – The Theory-Theory

The prototype and exemplars theories enjoyed a good deal of success during the late 60s and throughout the 70s. However, by the early 80s it had started to become clear that the prototype and exemplar theories, while accounting for a wide variety of experimental findings, were not adequate in themselves. Partly this was due to criticisms by philosophers, who argued (rightly) that people often are sensitive to considerations other than typicality and superficial similarity when categorizing and reasoning about the world.¹²² Partly it was due to an increasing number of experimental results which prototype and exemplar theories were not well equipped to explain. Both criticisms were based upon the same failing had by these theories: that prototypes and exemplars do not encode information of the kind needed to account for the full range of human cognitive behavior.

Prototype and exemplar theories were largely the product of work done in cognitive psychology. The theory-theory, by contrast, largely emerged from developmental psychology. As such, much of the early work in the theory-theory tradition was focused on the

¹²² Rey (1983, 1985).

categorization behavior of young children and how that behavior evolves over the course of their development.¹²³ Of particular interest was a clear shift in how children categorize that first emerges around roughly five years of age. Prior to age five, children seem to rely primarily on the superficial properties of things when categorizing them. For example, if a child is told a story where a cat is painted to look like a skunk, a child below age five will judge that the unfortunate animal is no longer a cat but now a skunk. A mere shift in its superficial appearance is sufficient in the child's mind to bring about a total change in its species. Relatedly, if told a story where a cow is opened up to reveal that it is full of gears, pulleys, and assorted mechanical devices, the child doesn't judge that this shows it is not a cow. Rather, it is taken to be a cow by the child, though perhaps a rather odd one. The faux cow's superficial similarity to real cows is sufficient in the child's mind to make it a cow. So far, this is exactly what the prototype and exemplar theories would predict: the children are making their categorization judgments on the basis of superficial similarities to typical category members.

Around the age of five this categorization behavior shifts. Children cease making categorizations that are so rigidly based on superficial appearances. Instead, they develop at least some sense that what makes many things the kinds of things they are is some deep property, something hidden from casual observation. They also develop a sense that these essential, hidden properties are what cause or account for the more superficial properties available to perception.¹²⁴ Past this age, the cat is now judged to remain a cat post-painting,

¹²³ Foundational works in the theory-theory tradition include Murphy & Medin (1985), Carey (1985), Keil (1989), and Medin & Ortony (1989). See Gelman (2003) for a more contemporary statement of the view.

¹²⁴ The view that people represent things as having hidden essences that explain their superficial properties is commonly called "psychological essentialism." See Medin & Ortony (1989); Gelman, Coley, & Gottfried (1994) for overviews of this position.

and the faux cow filled with mechanical innards judged to never be a cow at all. Accounting for these shifts in categorization behavior, as well as accounting for the categorization behavior of adults that failed to match the predictions of similarity-based theories, formed much of the impetus behind the theory-theory.

There are five central ideas that define the theory-theory approach to concepts:

1. Our knowledge of the world is organized into domains governed by domain-specific principles.
2. Our concepts encode information about the casual, functional, and/or explanatory relations between the properties of world-aspects.
3. Our concepts represent not just superficial but “deep” or “hidden” properties of world-aspects.
4. Our reasoning about the world is in some (often unspecified) manner analogous to reasoning in the sciences.
5. Our knowledge of the world develops over time in a way analogous to the way that scientific theories develop over time.

Regarding (1), according to the theory-theory we possess folk theories for domains such as physical objects, minds, artifacts, social relations, and biological kinds, with various other more niche theories proposed on occasion. And it is these theories that we rely upon these when categorizing and reasoning about things which fall within each respective domain. This is why, for instance, children from age five onwards refuse to categorize painted cats as skunks; around that age, their categorization of animals begins to be governed by a principle in their theory of the biological domain which states that animals are the kinds they are because of either their ancestry or some deep immutable property internal to them and hidden from

view.¹²⁵ To take another example, our categorizing and reasoning about the domain of artifacts is governed by a principle stating that artifacts have their functions essentially and that these functions are determined by the intentions of their creators.¹²⁶

Ideas (2) and (3) are intimately related, and as such I will treat them together. One issue with most (though perhaps not all) versions of the prototype and exemplar theories is that they represent the properties of their represented world-aspects in a disconnected way. A prototype may represent birds as typically having wings and flying, but it is silent on the relation (if any) between these two properties. By contrast, the theories posited by theory-theorists represent explanatory relations between the properties of world-aspect within that domain, e.g. that the property of being able to fly had by birds is explained by their property of having wings. And this is problematic, because an understanding of the relation between properties such as this seems to have an impact on categorization behavior.¹²⁷ It may be that thinkers take heaviness to be equally typical of barbells and bricks. But only the heaviness of the former is tied to its function, and thus a lack of heaviness counts against something being a barbell in a way it does not count against something being a brick.

Another issue with many (though again, not all) prototype and exemplar theories is that they tend to focus on superficial, perceptually-available properties.¹²⁸ That a skunk is black with a white stripe down its back is the kind of property that is the stock-in-trade of the prototype and exemplar theories, hence why very young children who likely rely on such

¹²⁵ Gellman & Wellman (1991).

¹²⁶ Bloom (1996).

¹²⁷ Murphy & Medin (1985).

¹²⁸ Though as I will argue this is an asset and not a failing if you see prototypes and exemplars as just two kinds of concept among many. Oftentimes superficial properties are highly useful in categorization and reasoning, sometimes even more so than deep or hidden properties.

concepts are led astray in cases where cats are painted to look that way. And as seen in the examples above of children categorizing animals past the age of five, theories can represent the hidden essences of objects, whatever property or properties hidden from perception account for the persistence of objects across superficial changes, e.g. that property or properties of cats which accounts for their continuing to be cats both before and after their being painted.

The other defining ideas, (4) and (5), of the theory-theory I will not spend as much time on, as ultimately my concern is more with the structure of concepts under the theory-theory than with their attendant processes and methods of acquisition.¹²⁹ The first of these is that the theory-theory understands the processes involved in categorization, inference, and the like on the model of scientific reasoning. Whatever processes the child utilizes when categorizing are, in some usually unspecified manner, analogous to those used by scientists.¹³⁰ Second, the theory-theory understands conceptual development as being analogous to theory change in the sciences. In other words, changes in how a child thinks about the world over time (their shift, for instance, from a more similarity-based way of thinking to a more essentialist one) is understood as similar to how new scientific theories emerge and supplant earlier ones.¹³¹

From all this, it may be unclear what exactly a concept is under the theory-theory. This is because psychologists themselves are inconsistent on this point. Some psychologists hold

¹²⁹ Issues of conceptual development will, however, be of importance and returned to in ch. 4 and 5.

¹³⁰ This is often taken very literally. See for instance Gopnik's *The Scientist in the Crib* (2000). Though see Machery (2009, 101-102) for worries that theory-theorists are less working from an analogy with the practice of actual scientists, and instead working from a more folk, intuitive understanding of scientific activity.

¹³¹ Keil & Batterman (1984). However, Keil (1989, ch. 12) argues that this does not show a shift from children relying on properties taken to be inessential to those taken to be essential when categorizing, but rather a shift in which properties they take to be essential.

that concepts simply are theories, e.g. my concept of water just is my theory of water.¹³² Others hold that they are portions of larger domain-specific theories, e.g. my theory of water is that portion of my folk theory of chemical substances that bears on water.¹³³ Lastly, and most commonly it seems, some hold that concepts are the constituents of theories in a way analogous to how theoretical terms are constituents of scientific theories.¹³⁴ This last way of understanding concepts carries with it the advantage that it allows for the pairing of the theory-theory with a conceptual role semantics, and this is an advantage that some within the theory-theory tradition have recognized.¹³⁵ Carey, for instance, holds that concepts should be “analyzed relative to the theories in which they are embedded” and “identified with the roles they play in theories.”¹³⁶

There is comparable disagreement over what exactly counts as a theory in the sense that theory-theorists are concerned with. Some psychologists are permissive on this point, allowing for nearly any organized body of beliefs to count as a theory.¹³⁷ Others are highly restrictive, limiting theories to specific bodies of belief regarding a limiting range of domains and with the right organizational structure.¹³⁸

¹³² Rehder (2003), Rips (1995).

¹³³ Prinz (2002) often seems to understand concepts under the theory-theory in this way.

¹³⁴ See Gopnik & Meltzoff (1997), Keil (1989).

¹³⁵ It should be noted that concerns about semantics seem to be either totally ignored by psychologists in many cases or even outright dismissed. Carey (1985, 2009) is a rare exception to this.

¹³⁶ Carey (1985), pg. 198.

¹³⁷ Murphy & Medin (1985), pg. 436 lists five properties of theories: (1) they provide explanations of kind with regard to a given domain, (2) they simplify reality, (3) they have external structure, (4) they have internal structure, and (5) they interact with data and observations in some way. As stated, it does not seem that these properties will exclude many organized bodies of beliefs.

¹³⁸ Carey (1985), pg. 201: “My conjecture ... is that there are only a relatively few conceptual structures that embody deep explanatory notions—on the order of a dozen or so in the case of educated nonscientists. These conceptual structures correspond to domains that might be the

The theory-theory has much to recommend it. It explains facts about both the changes in ways of thinking observed across cognitive development as well as facts about our tendency towards essentialist thinking that similarity-based theories fail to account for. However, it cannot by itself be a fully adequate theory of concepts. This is because the information which is encoded in the theories, while useful as a constraint on categorization, is not comprehensive enough or of the right kind to make reliable categorization judgments. This is so even if one looks at categorization in context, focusing solely on realistic cases and excluding all odd, implausible, or overly philosophical possibilities. The issue is that the theory-theory shares the same problem as the prototype and exemplar theories but in the opposite direction: prototypes and exemplars cannot explain categorization that relies on information that goes beyond the superficial, while the theory-theory fails to explain cases where categorization really does seem to rely mostly on those superficial characteristics.

Why this is so can be understood by looking at the usefulness of prototypes and exemplars. Consider a simple example like that of categorizing a dog as such. Much of why prototypes and exemplars are useful for this task is that they represent properties that are easily identified in perception or through a small number of easily made inferences. I can simply tell at a glance that a given dog has the typical properties represented by my prototype or exemplar set. I can just see that it is four-legged, has a body of a certain shape, is covered in fur. I can hear its bark. I may even smell one of the many typical canine odors. Those typical properties that are not directly observable, such as one's pertaining to the dog's behavior, can be inferred from properties that can be directly perceived, such as the movement of its body and how it

disciplines in a university: psychology, mechanics, a theory of matter, economics, religion, government, biology, history, etc.”

interacts with things in its immediate environment. I can, for instance, recognize that it wags its tail by first recognizing its tale and then observing its movement.

While it's true that not all dogs will have these properties, especially those dogs which only inhabit merely possible worlds, in context this is often of little to no practical relevance. I rarely, if ever, have occasion or need to categorize as dog or non-dog such entities as swampdogs or dogs altered in radical ways via genetic and physiological tinkering. The overwhelming majority of dogs I encounter are mundane, typical, bog standard dogs. Some may deviate from what is typical in a few minor respects. I may encounter a three-legged dog, a hairless dog, or a dog unable to bark due to being debarked by its owners. However, these deviations are usually not enough for my dog prototype to fail to generate a positive categorization judgment. Prototype theory does not demand total overlap in typical properties, only that there is overlap to a sufficiently high degree. When there is no such overlap, such as when the differences occur together thus rendering the dog insufficiently similar to my prototype, the prototype and exemplar theories correctly predict people's real tendency to either waver or outright fail to make the correct categorization judgments. A dog sufficiently diseased, such as one suffering from severe mange, may not be recognized by most people as being a dog, and this is entirely because of the fact that such a dog bears little similarity to most people's idea of a typical dog.

What makes prototypes and exemplars useful, in short, is that they represent properties that are highly available to thinkers. They represent properties either immediately available in perception or which require relatively little investigation or thought to discover. The theory-theory by contrast focuses on properties of kinds that are by comparison not available or discoverable in any practical sense to most people. I cannot determine a dog's phylogeny in

the same immediate way I can immediately determine its furriness. My theory of a kind may represent its essential, causal, or functional properties. But that information is often useless for categorization purposes unless I can recognize the presence of those properties when I encounter them in the world in the sorts of contexts I actually find myself in.

Worse is that situations philosophers may be inclined to imagine when considering the theory-theory in the abstract, situations where individuals know and can draw upon their knowledge of the deep properties of things in a variety of cognitive tasks, is the best-case scenario for the theory-theory. The number of people throughout human history who have had knowledge of, for example, the essential properties of the things they interact with daily is vanishingly small. For the majority, the most they are able to attribute to things is a so-called placeholder essence.¹³⁹ They know that there is some *something*, some deep or hidden property of a thing that both makes it the kind of thing that it is and which is the cause of many of its more superficial properties. They know, in other words, that the thing has an essence of some sort. They cannot, however, say much of anything (if anything at all) about just what this essence is, and in many cases simply guess based on their limited and fragmentary knowledge. This knowledge is nowhere near robust enough to make reliable judgments.

When the five-year-old child refuses to categorize the unfortunate cat striped with paint as a skunk, it is unlikely to be because they check for the presence of feline essential properties

¹³⁹ Medin & Ortony (1989, pg. 184). describe psychological essentialism as “not the view that things have essences, but rather the view that people’s representations of things might reflect such a belief (erroneous as it may be) ... we propose that the knowledge representations people have for concepts may contain what might be called an *essence placeholder*. In some cases ... it might be filled with beliefs about what properties are necessary and sufficient for the thing to be what it is. In other cases it might be filled with a more complex, and possibly more inchoate, ‘theory’ of what makes the thing the thing that it is.”

both before and after the painting, finding them present in both instances. Nor do they likely check for skunky essential properties post-painting. They categorize as they do because the child (1) knows that the animal prior to the painting was a cat (something explicitly told to them by those running the experiment), (2) that what makes something a cat is not its superficial properties but some properties inaccessible to casual observation, and (3) that painting a cat is a superficial change to the cat. The child's folk biology is sufficient to give them knowledge of the kind in (2) and (3). It cannot though give them knowledge of the kind in (1). The child needs to first know that something is a cat before they can judge that it remains a cat through some superficial change. That kind of knowledge is likely only securable by the child through their understanding of the superficial properties of cats, i.e. through the use by them of prototypes and exemplars, at least outside of the artificial confines of an experimenter's workspace. Thus, a full explanation for the child's categorization behavior cannot dispense with prototypes and exemplars. Similar points apply to the categorization behavior of adults and to the other core cognitive capacities. I categorize a substance as water not because I recognize it as having the chemical structure H_2O , but rather infer that it has that structure based upon my categorization of it as water based upon my awareness of its superficial properties. That water is H_2O is a constraint upon, rather than the means by which, I categorize some substance as water in realistic contexts.

Were it the case that we both knew enough about the world and could also immediately or easily discern the deep properties of things, then theories alone could be sufficient. We do not though know enough and cannot easily or quickly determine something's deep properties. In the real world we work with much more limited information, and that information is not always easy to come by. We see, smell, taste, touch, and listen to things. We poke and prod

them, sometimes with a long stick and other times with sophisticated scientific instruments. We often do so under tight time pressures that do not allow for lengthy investigations. And there are consequences if we are wrong or do not make our judgments quickly, sometimes minor and sometimes life-ending. Had we eyes like microscopes and fingers like strips of litmus paper, or if we lived in a world where the essential properties of things bulged under the skin or were stamped across thing's surfaces, then things would be different. As we do not, we have no choice but to often rely upon superficial, inessential properties when determining what is what in the world.

2.4 – Other Theories

While the classical, similarity, and theory-based approaches to concepts are the dominant ones in the psychological literature, they by no means exhaust all the theories on offer. The remaining theories I will consider should not, I think, be thought of as competitors to the theories considered earlier. Rather, they should be thought of as supplemental and complementary. Some of them are representations that encode different kinds of information than definitions, prototypes, exemplars, and theories, and thus are suitable for different ends and different contexts. These are neoclassical, ideal, and hybrid concepts. Others differ, not in what they represent, but in how they are assembled, used, and discarded. These are the so-called ad hoc concepts. I will briefly address each of these in order to give a sense of what other possible forms concepts may take.

2.4.1 – Neoclassical Concepts

Neoclassical concepts are meant to address one of the main failings of the classical theory, namely the seeming lack of definitions. Because of this problem, neoclassical concepts do not purport to represent necessary and sufficient conditions for category membership.

Rather, they encode only partial definitions of what they represent, i.e. they solely represent necessary, but not sufficient, conditions for category membership.¹⁴⁰

Take something like the concept of knowledge characteristic of the JTB account. A classical concept representing knowledge as justified true belief would be inadequate for the myriad reasons epistemologists have raised in recent decades. It could, however, be a perfectly adequate neoclassical concept to the extent that one thinks that being a belief, being justified, and being true are still necessary conditions on knowledge. Whatever else may be needed for sufficiency (if any fourth condition is even possible) would not need to be part of the concept.

While the definitional properties of things seem to be unavailable to ordinary (and even expert) thinkers in nearly all cases, the necessary properties of things are comparably easier to grasp. I have no idea what fourth condition, if any, would render a justified true belief knowledge, but I am at least confident that knowledge must be at minimum true. I believe (perhaps wrongly) that rock music must at least have a guitar somewhere in the mix, but I have not the foggiest idea how to separate it from genres, like some harsher forms of folk and blues, that butt up against the category and also make use of that instrument. Knowing (or at least believing) that things have necessary properties is common. Thus it should not be a surprise that we have representations which support these beliefs, and neoclassical concepts seem to be a plausible candidate.

¹⁴⁰ See Jackendoff (1983) and Pinker (1989) for two characteristic statements of the view. Note, however, that each of their views may be more accurately considered hybrid views, as the concepts they posit include information beyond partial definitions, e.g. Jackendoff's concepts include not just partial definitions, but include (among other things) information about typical properties as well.

2.4.2 – Ideals

Ideals are another less commonly encountered kind of concept. Ideals are, in many ways, similar to prototypes and exemplars.¹⁴¹ Like them, ideals represent the typical or cue-valid properties of a kind. Further, they also factor in similarity-matching processes like those found in prototype and exemplar theories. Ideals represent, however, not typical members of a kind, but rather ideal members.¹⁴² In other words, they represent not what is typical of members of a kind simpliciter, but rather what is typical of those members taken by the concept possessor as ideal for that kind.¹⁴³

Consider the following example. If I am on a date, I presumably have as my goal figuring out whether the other person would be a good romantic and/or sexual match for me. In determining whether or not this is the case, I certainly should not be concerned with whether or not this person is merely similar to the typical partner I have had in the past or to some summary or average of past partners for both myself and others. This is doubly so if I have poor taste or a poor track record when it comes to dating, and thus in this instance want anything but what is typical. What I should care about is how similar this person is to some idealized partner. I should care how many of the properties of what I take to be good in a

¹⁴¹ This is not a coincidence, as the theory of ideals emerged from the similarity-based tradition that includes prototype and exemplar theories. See Barsalou (1983) for the original statement of the theory of ideals.

¹⁴² It could be argued that prototype and exemplar theories can accommodate ideals. I may, for example, have a prototype of IDEAL PARTNER as opposed to simply having a prototype of PARTNER. This is, however, very likely to be an empirical question.

¹⁴³ What is represented as being ideal or valuable is, presumably, relative to the concept possessor's desires and goals, e.g. my ideal CAR will likely include very different features if I need a car to take me to and from work versus needing it to shuttle not just myself but an entire family around town or to haul heavy loads.

partner they possess and how much each property should be weighted in my post-date deliberations.

One thing worth noting about ideals is just how common they seem to be. When considering what things to sell at a garage sale (lacking emotional attachment, having broad appeal), choosing a parking spot at the supermarket (far from the entrance and far from the cart wells), or making dinner for the night (cheap and with abundant opportunity for leftovers), I am being guided by concepts derived from my various goals and which represent the ideal that I am aiming for in my actions. Thus, while perhaps less discussed than concept types like prototypes and exemplars, this is not in any way due to their marginal status in our cognitive and practical lives.

2.4.3 – Ad hoc Concepts

Importantly, ideals are often thought to differ from more traditional concepts in one very important and central respect. Concepts are commonly thought of as something that once acquired are possessed in perpetuity, save for injury, disease, death, or perhaps even simple forgetfulness. This is so common that it is often built even into the definition of concepts. To take one example, Machery defines concepts as follows:

A concept of x is a body of knowledge about x that is *stored in long-term memory* and that is used by default in the processes underlying most, if not all, higher cognitive competences when these processes result in judgments about x .¹⁴⁴

While this definition has much to recommend it, one major issue is that it rules out the possibility of ad hoc concepts by fiat. Ad hoc concepts are concepts which are “not established

¹⁴⁴ Machery (2009, pg. 12).

in memory but derived impromptu to achieve a current and novel goal.”¹⁴⁵ In other words, they are concepts which are assembled on the fly for use in a single task or for a relatively short period of time and stored in working memory for that duration. They are not (or at least need not) be stored in long-term memory after their usefulness has ended.

One of the classic examples of an ad hoc concept is due to Barsalou., which I have adapted to a form that should be more familiar to many Californians in recent years¹⁴⁶ Suppose I wake up in the middle of the night to an alert on my phone. Upon checking it I learn that a nearby wildfire is threatening my home and I am now under a mandatory evacuation order. I have a very limited amount of time to determine what, if anything, I can take with me as I escape. To guide my decisions I construct an ideal, THINGS TO TAKE FROM ONE’S HOME DURING A FIRE that represents the properties typical of something in that category, e.g. VALUABLE, IRREPLACEABLE, and so on. I then use this ideal to guide my decisions as to what to take with me, comparing candidate objects against the ideal and looking for similarity in their respective properties. Once I have collected my things and am somewhere safe, I no longer have any use for the ideal I constructed, and I thus purge it from my short-term memory.¹⁴⁷

That is not to say that ad hoc concepts are limited to ideals, however.¹⁴⁸ In many cases, prototypes, exemplars, and perhaps even definitions can be ad hoc. For instance, upon parking I may create an exemplar of the specific space in which I parked or a cognitive map which I construct as I walk to my destination, and then use that to guide my finding my way back to

¹⁴⁵ Barsalou (1991).

¹⁴⁶ Barsalou (1983).

¹⁴⁷ Though of course I may eventually commit it to long-term memory if I, say, live in California and this is a regular enough occurrence to warrant my doing so.

¹⁴⁸ Nor are all ideals ad hoc, e.g. my concept of the ideal boyfriend is likely to be something I keep long-term, and likely update over time in light of my experiences.

that space on my way out. Or while grading a large stack of papers I may construct a prototype of what a typical paper looks like, then grade papers up or down based on how much and in what way they deviate from what is typical.¹⁴⁹ When conditions are such that the purpose for which a concept is suited is a purpose that is novel and unlikely to reoccur, that concept is likely to be ad hoc in nature. There is little point in me storing exemplars for each and every parking space I have ever parked in, and likely much potential for confusion were I to do so.

Like ideals, ad hoc concepts receive comparatively less attention than their counterparts stored in long-term memory. And like ideals, this comparative lack of attention is likely not due to their being uncommon or in any way marginal. Some theories of concepts even hold that all concepts are ad hoc, perhaps being assembled on the basis of background knowledge for particular purposes before being discarded, or perhaps it is rather the whole of cognition which involves fleeting, ad hoc representations.¹⁵⁰ I draw attention to them here to ensure that the reader is aware of their possibility, and thus aware of the expansion of possible objects of

¹⁴⁹ I of course could (and perhaps even should) use an ideal for this purpose, developing a sense of what I think the paper ideally should be and then grading based on how far the paper deviates from this ideal. But there is no impossibility in using a prototype, say for a B paper, for this purpose instead (it is, in fact, how I personally tend to grade).

¹⁵⁰ See Casasanto & Lupyan (2015) for an argument in favor of this latter position. Barsalou (1987) also appears to be sympathetic to this position: “Knowledge in long-term memory ... is relatively undifferentiated and continuous. ... It may be extremely difficult, if not impossible, to identify where the knowledge for a particular category in long-term memory begins and ends. To the extent that this is true, it is hard to imagine how there could be invariant representations for categories stored in long-term memory.” While this may be possible under certain paradigms (Casasanto and Lupyan appear sympathetic to embodied and enactive approaches), I do not see how it could be workable with the assumptions I am working under. Ad hoc concepts would need to be constructed on some basis, and all I can think of that could serve this basis is either background or conceptual knowledge. It could not, I suspect, be ad hoc concepts all the way down. As such, I am committed to there being at least some concepts (and likely the majority) that are stored in long-term memory.

conceptual representation, the cognitive flexibility concepts in isolation can afford, and the diverse contexts of use and purposes towards which they can be put.

2.4.4 – Hybrid Theories and Concept Pluralism

Finally, there is the hybrid theory and concept pluralism. Much like the earlier theories, neither the hybrid theory nor concept pluralism are really individual theories in the strict sense. Rather they are more broad families of theories, though ones far broader and more open-ended than those earlier discussed. As hybrid and pluralist theories form much of the focus of the next chapter, my discussion here will be comparatively brief.

On hybrid theories concepts divide into parts, each of which encodes a specific kind of information characteristic of one of the concept kinds discussed earlier. For instance, one classic version of the hybrid theory holds that concepts divide into a conceptual core that encodes a definition and a prototype that forms the concept's periphery. Osherson and Smith give a statement of one classic version of the hybrid theory:

We can distinguish between a concept's core and its identification procedure; the core is concerned with those aspects of a concept that explicate its relation to other concepts, and to thoughts, while the identification procedure specifies the kind of information used to make rapid decisions about membership . . . We can illustrate with the concept woman. Its core might contain information about the presence of a reproductive system, while its identification procedure might contain information about body shape, hair length, and voice pitch. Given this distinction, it is possible that some traditional theory

of concepts correctly characterize the core, whereas prototype theory characterizes an important identification procedure.¹⁵¹

On this view, when categorizing one first checks the perceived object against the prototype. If the degree of match is high enough, there are tight time pressures, or low stakes, then a positive categorization judgment is made. But if the similarity is borderline, there is ample time, or the stakes are high enough, the core (such as a definition) is checked as well. In cases where the definition fails to apply, the similarity to the prototype is disregarded and a negative categorization judgment made instead. If core and identification procedure agree, then the positive judgment is maintained. Hybrid concepts need not consist of definitional cores and prototypical peripheries. Hybrid concepts can come in many forms, depending on how many and what kinds of parts they take concepts to divide into.

By contrast, concept pluralism in its simplest form is the view that concepts are heterogeneous, usually with regards to their coarse structures.¹⁵² In other words, it is commonly understood as the view that prototypes, exemplars, theories, and the like all exist and count as concepts. Like the hybrid view, conceptual pluralism comes in a wide variety of forms depending on which concept kinds one includes. In the following chapter, my goal will be to argue that concept pluralism, properly understood, is the theory most likely to provide a plausible account of what both the scientific literature and our own common sense seem to show about concepts. Further, I will argue that the necessary form of concept pluralism is more robust than often thought and allows for a very wide range of concepts differing in terms of both coarse and fine structure.

¹⁵¹ Osherson & Smith (1981).

¹⁵² See Machery (2009, ch. 3) and Weiskopf (2009a, 2009b) for defenses of concept pluralism.

Theory	Structure	Information Encoded	Process
Imagism	Pictorial representation	Perceptually available properties	Similarity matching
Classical	Representation of a definition	Necessary & sufficient conditions	Rule-based
Neoclassical	Representation of a partial definition	Necessary conditions	Rule-based
Prototype	Representation of a typical/representative kind member	Typical/cue-valid properties	Similarity matching
Exemplar	Representation of particular or set of such representations	Typical/cue-valid properties	Similarity matching
Ideal	Representation of an ideal category member	Ideal properties	Similarity matching
Theory-Theory	Domain specific theories	Causal, function, & historical properties (among others)	Theoretical reasoning
Hybrid	Structured complex of two or more kind of concept	Structured complex of the information encoded in two or more kind of concept	Two-step: similarity matching followed by overriding rule checking
Pluralism	Variable	Variable	Variable
Atomism	None*	None*	?
Eliminativism	n/a	n/a	n/a

Figure. 2.5

3 – Conceptual Pluralism

In the previous chapter I surveyed the leading theories of concepts from cognitive and developmental psychology. Implicit in that discussion were two major points relevant to the development of any theory of concepts. First, is that none of the current theories on offer is likely to ever be a fully adequate theory of concepts or cognition in general. Second, is that there is no reason why one could not hold that each of the kinds of concept posited by psychologists, along with their corresponding cognitive processes, exist. What I propose to do in this chapter is draw out some of the implications of these points and argue from them to a robust form of conceptual pluralism, a position where each theory captures but one kind of concept and its attendant processes. Such a pluralism is robust in that it is not just a pluralism of coarse conceptual structure but of fine conceptual structure as well. Thus, it is not just that concepts are comprised of a wide variety of representational kinds, such as prototypes, exemplars, and the like. In addition, there is a great deal of heterogeneity between the particular concepts falling within each of these classes.

In addition, in this chapter I aim to establish in addition two supporting, though still important in isolation, positions on concepts: (1) that the background/conceptual knowledge distinction is, as a matter of empirical fact, but not metaphysical necessity, either nonexistent or otherwise without explanatory relevance, and (2) that hybrid and pluralist theories of concepts need not be competing theories but can instead plausibly complement one another.

3.1 – Coarse Conceptual Pluralism

As briefly argued in chapter 2, no one theory of concepts is likely to adequately account for the full range of cognitive resources drawn upon in categorization and related tasks. Each of the theories considered alone struggles with cases of some kind and each seems to leave out

something important with regards to cognition as a whole. This is something that only becomes more clear the longer one looks at the experimental findings. For every finding that seems to support the existence of prototypes, to take one example, there is another that they fail to account for. Prototypes do well in explaining quick categorization judgments or those of very young children, but they struggle to explain the more considered judgments of older children and adults. One sees similar things for exemplars, theories, and the various other concepts posited by psychologists, with each explaining why some judgments are made while struggling to explain others. Further, it is not clear how these theories could be supplemented to avoid this, particularly given that concepts of each kind represent a fairly limited sort of information, e.g. prototypes representing only typical and/or cue-valid properties of objects, or exemplars representing only particulars and their properties.

Even if one could build into these representations a kind of information that they do not standardly include, there would still be the problem of the processes corresponding to these conceptual kinds. One could, for instance, insist that prototypes can encode information about the deep properties of things, properties like something's function or its hidden essence. Or perhaps the prototype can represent in addition the causal relations between the typical properties represented by the prototype, e.g. that FLYING and HAS WINGS for BIRD are causally related insofar as it is the having of wings which allows for birds to fly.¹⁵³ But the prototype would still be representing these properties as merely typical, rather than essential or necessary. And they would still be factoring in similarity-matching processes that result in graded judgments, rather than the often binary rule-based judgments characteristic of definitions and theories. Such attempts to supplement one of the existing theories will still leave a gap to be

¹⁵³ Hampton (2006), Malt (2010).

filled in, some kind of information or cognitive process which we do in fact use to categorize and reason with, but which the kind of representation and process posited by the theory fails to accommodate.

Further, while there is no shortage of debate between the proponents of the above theories, they need not be treated as competing theories at all. Each of the kinds of concept posited by the major theories encodes a different sort of information, has distinctive processes associated with it, and is useful in its own way depending on its context of use and the ends of its user. When doing geometry problems in school or charting out your family history, you are going to be best served by definitions and the consideration of the strict necessary and sufficient conditions that they encode. This is because the domains in question are sufficiently orderly and the definitions widely known enough to make this feasible.¹⁵⁴ If instead one instead needs to quickly recognize a student in lecture, then rather than looking for a set of (likely unavailable) necessary and sufficient conditions for being that student, one ought to be looking for similarities to an exemplar representing such easily recognized properties as their typical dress, hair, makeup, and the like. When predicting the behavior of another intelligent creature, reliance on a folk theory of mind is most likely to bring success, and it's difficult to see how prototypes, exemplars, or definitions could even hope to begin to provide the basis for attributing beliefs and other propositional attitudes. If someone is picking out a new television or making some other major purchase involving extensive deliberation, they will likely do best comparing possible purchases against some represented ideal. And so on.

¹⁵⁴ A prototype in these cases may, for instance, lead one astray by latching onto merely contingent properties of well-defined world-aspects. A prototype UNCLE, representing properties such as *tells lame jokes at family reunions*, is unlikely to be much help in the task of constructing a genealogy of one's extended family.

There is no reason to think, so far as I can see, that a creature with a cognitive and practical life as diverse and complex as that of humans would not draw upon an equally diverse set of cognitive tools.¹⁵⁵ Nor can I see much reason to think that the even greater diversity found between the many thinking species we know of, with their wildly different environments and forms of life, can be accommodated by any single kind of coarsely structured concept and its corresponding processes.

Neither the inadequacy of individual theories nor their compatibility with one another should, I think, be surprising, as both points comport with common sense. We do seem to use information about prototypical category members, sets of category members, loose theories involving these categories, ideals of what such category members should be like, and so on all the time, sometimes all even in service of the same task. We rely upon a wide variety of information and reasoning strategies in navigating our day-to-day lives and pursuing our various ends. Why would there not then be representations and processes corresponding to, or at least including, each of these sorts of information and processes corresponding to these kinds of strategies? None of the theories surveyed was expansive enough to be singularly and fully adequate as a theory of what makes cognitive capacities such as these possible, and there was never much reason to suppose beyond a bare appeal to the value of parsimony that any one of these theories ever needed to carry such a heavy explanatory burden.

Considerations such as these, along with the fact that half a century of debate has yet to result in anything approaching a consensus position in psychology, give at least some reason

¹⁵⁵ I similarly doubt that the cognition of many non-human animal species (considered individually), could be accounted for in this way either. This is so even for non-human animals often considered “lower.” Though I will not argue for that here, see ch. 7 for some consideration of these issues.

to consider the merits of more pluralist options. If the kinds of information drawn upon by core cognitive capacities outstrips the resources provided by any one kind of concept, why not then suppose that each of these concepts is real and drawn upon in different contexts and for different tasks? Why not, for instance, suppose that prototypes are relied upon when making quick categorization judgments of ordinary perceptible objects under tight time constraints, the exact contexts and tasks where prototype theory has had the most explanatory success? And why not suppose that theories are utilized in cases where there is time and reason for more considered and careful judgments, cases where the theory-theory has found much success, but where prototypes, exemplars, and other theories of similarity-based concepts have long struggled? And lastly, why not suppose that definitions are in fact used in the minority of cases where definitions seem to actually be available and may be uniquely well-suited to the task at hand? There seems to be little gained beyond parsimony (and even then, only parsimony of an ontological sort) and much to be lost explanatorily for insisting otherwise, for insisting that we must have one simple theory of concepts to explain all of cognition.

There seem to be good reasons then to suppose that there is a great deal of heterogeneity when it comes to the coarse structures of concepts. We then ought to, I propose, allow a wide variety of concept kinds into our ontology, however many are needed to do the explanatory work expected of concepts. Definitions, prototypes, exemplars, theories, and ideals all exist as distinct concept kinds. Distinctive processes exist that make use of each of these kinds of concept, such as similarity-matching processes for prototypes, exemplars, and ideals, rule-based processes for definitions, and processes approximating scientific theorizing for theories. And each of these kinds of concept are deployed from long-term memory or assembled on the

fly in context and task-sensitive ways, either individually or in coordination with one another to achieve the thinker's ends.¹⁵⁶

There are a few clarifications worth making prior to continuing any further. First is that the conceptual pluralist is in no way committed to prototypes, exemplars, theories, and the other coarsely structured concept kinds being the only such concept kinds there are. They are also not committed to these being the only *possible* coarsely structured concept kinds, even if they turn out to be the only such concept kinds in actuality.¹⁵⁷ The conceptual pluralist can allow, and should be willing to, supplement the current crop of concept kinds as new psychological evidence comes in. In fact, there is no shortage of potential candidates. Perhaps, for instance, perceptual representations and processes are sometimes used in categorization and inferential tasks, as in the theories of concepts endorsed by neo-empiricists like Prinz and Barsalou.¹⁵⁸ Cognitive maps may be another case of a possible conceptual representation.¹⁵⁹ Lastly, cognitive semantics offers a wide variety of possible concept kinds, such as image schemata, that may end up being needed for a fully adequate theory of linguistic comprehension.¹⁶⁰ This is not to endorse any of these views or the claim that they are concepts in the sense relevant here, but rather to merely note that there is nothing inconsistent in the conceptual pluralist including them in their ontology of concepts. And of course, there is nothing inconsistent about the conceptual pluralist dropping one of these concept kinds, or

¹⁵⁶ Pluralism of concepts in terms of their coarse structures has been defended by Machery (2009) and Weiskopf (2009a, 2009b).

¹⁵⁷ If artificial psychologies are possible, then there is presumably no limit to the number of possible coarsely structured concept kinds. The coarse structure of concepts would then be limited only by what designs human minds could come up with and implement.

¹⁵⁸ Barsalou (1999), Damasio (1989), Glenberg (1997), and Prinz (2002).

¹⁵⁹ See Camp (2007) that such cognitive maps can be used to think in much the same way as proposed linguistically-formatted mental representations such as formulas of mentalese.

¹⁶⁰ Evans, Green (2006).

even more traditional concept kinds like prototypes and exemplars, if it is shown they are not necessary for explanatory purposes.

Concept pluralism also need not imply that everything can be represented by every kinds of concept. It may be that some things can only be represented by concepts of certain kinds, or that while some things can be represented by multiple concept kinds as a matter of empirical fact, we only utilize one concept type to represent it. For instance, it may be that particulars are only able to be represented by exemplars, or that the hidden or deep properties of things can only be represented in terms of theories. What, after all, would a prototype of an animal's lineage or essence include as its features? Or what would an exemplar of a quark represent, if individual quarks can even be represented in human thought?

The pluralism I am arguing for here, however, concerns not just the coarse structures of concepts and their corresponding kinds, as we have been discussing up until now. Rather, it includes in addition a commitment to a pluralism of concept kinds defined by their fine structures, a commitment to not just, for instance, prototypes and exemplars of cats, but a wide variety of prototype and exemplar kinds that represent cats in differing ways. In the section that follows, I will defend pluralism of this sort.

3.2 – Fine Conceptual Pluralism

As discussed in chapter 1, there is a tendency in philosophy to think that, while there may be some variation between the concepts we possess, this variation is reasonably bounded. There may be something like the folk concept of W and the concept of W had by experts, or one culture's concept of W versus the concept of W had by members of another culture. But the variation between concepts, if any, is nowhere close to being as pervasive and severe as psychologists and concept pluralists are often inclined to think.

This is, I think, a mistake. Concepts are highly variable, and in more than one way. As already discussed, concepts vary in terms of their coarse structures, with concepts of a given thing being implemented by prototypes, exemplars, theories, ideals, and so on, each of these often occurring together within the mind of a single thinker. Further, as I will now argue, concepts vary greatly in terms of their fine structures, the specific ways in which particular prototypes, exemplars, and the like represent their referents as being. In addition, such variability in fine structure is both inter and intrapersonal: concepts vary in their fine structures both between individual thinkers and across stages of a single thinker's cognitive development.

There is extensive evidence of such variation in the psychological literature, and in fact such variation seems to more often than not to be simply assumed.¹⁶¹ For instance, while subjects' typicality judgments are fairly consistent when it comes to highly typical category members and those that are entirely unrelated to the category, those members in the middle tend to produce inconsistent judgments.¹⁶² Given that typicality judgments are thought to reflect the fine structure of prototypes, this gives reason to believe that different subjects have prototypes that represent the category differently, likely representing either different properties as typical of the category or differing in the weights they assign to these. Further, it is common for subjects to outright disagree about category membership for marginal category members, and this is thought to again reflect differences in either what properties are represented or how they are weighted. For instance, subjects disagree about whether activities like darts and hiking

¹⁶¹ See Hampton (2020) and Murphy (2020) for overviews of some of this evidence.

¹⁶² McCloskey & Glucksberg (1978).

count as sports, and this appears to reflect differences in how they weight properties like physical exertion versus the need for skill in the activity.¹⁶³

Concepts are highly variable in their fine structures in large part because their fine structure is a function of a variety of idiosyncratic properties of their possessors. What fine structures their concepts possess and thus what cognitive content those concepts encode depends on, among potentially many other things:

1. Their possessor's distinctive course of experience.
2. Their possessor's stage of cognitive development.
3. Their possessor's cognitive, perceptual, and practical capacities.
4. The context in which the concept is acquired or is to be used.
5. The purposes towards which the concept is to be put.

Differences along any of these dimensions will, in the normal case, result in structural differences between the relevant conceptual representations.¹⁶⁴ I will go through each of these dimensions in order, saying a little about how and why they affect the fine structure of concepts.

3.2.1 – Differences in Experience

Consider first differences in formative experience. Concepts like prototypes and exemplars reflect what is typical, not of the category as a whole, but rather what is typical of those members of the category the concept possessor has themselves encountered, whether in person or via representations on television, in books, and the like. And these differences in

¹⁶³ Verheyen & Storms (2013). See Zee et. al (2014) for similar findings.

¹⁶⁴ Note that I do not here use “normal” to mean typical. By “normal” I mean something like the sense in which hearts normally pump blood or beavers normally build dams. When I mean typical specifically, I will always use the word “typical” instead of “normal.”

what is encountered will likewise be reflected in differences in fine conceptual structure. Someone whose main exposure to birds has been ostriches and emus is unlikely to have a prototype of birds that is structurally similar to the prototype of someone who mainly encounters pigeons and seagulls.¹⁶⁵ The latter's prototype is likely to including FLYING as one of its constituents, but the former's is not. The former's is likely to include AGGRESSIVE or TALL as constituents, while the latter's likely will not. And given the kinds of birds these people encounter and have practical dealings with, it is to their benefit to have the specific prototypes they have. One would not want to engage with emus, with their six inch talons, in the same manner that they engage with pigeons and seagulls.

Some concept types are even more variable due to changes in experience. Finely individuated exemplars may even be in principle private and unsharable, or at least private as a matter of practical reality if not metaphysical necessity. The set of exemplars that one uses in tasks such as categorization will reflect one's own idiosyncratic experience. My exemplars for cats, for instance, will include exemplars for past pets of mine (Butterscotch), or cats with which I've had sufficient interactions such that I have a need to commit information about them to long-term memory (Rylee). Your exemplars will likely be quite different, not including exemplars for such cats as Rylee, Buttercup, and Mowman, but rather ones for your own pets or neighborhood strays.

¹⁶⁵ This does depend on understanding "encounter" in a particular way. Someone from Australia may see more emus in person than they see seagulls, but they are likely to see far more seagulls through media than emus. The point still stands either way. What members of a kind are judged to be typical will be influenced by one's experience with that category. All other things being equal, greater exposure to and experience of certain members of a category will increase how typical one finds those members relative to those less experienced.

Even if we are exposed to the same particulars and both form exemplars of them, these exemplars will reflect the times of our exposure to those particulars and those properties most salient to us. A young child's exemplar of their mother will, one assumes, look quite different from the exemplar had by the mother's romantic partner, her boss, her own parents, or even her pet dog. Each represents quite different features and assigns differing importance to those features that may be represented in common. A child may plausibly and reasonably represent their mother via an exemplar that includes features like FOOD PROVIDER, but one hopes her partner and boss do not do the same.

3.2.2 – Differences in Stage of Cognitive Development

Regarding (2), concepts vary in terms of both their coarse and fine structures across the stages of a single person, most notably during early childhood development. The development of theories across childhood is the clearest example of this. As discussed in chapter 2, the child of five who recently ceased categorizing storybook cows filled with pulleys and gears as cows has undergone a shift in the structure of their concept of cows, or perhaps the acquisition of a new cow concept altogether.¹⁶⁶

Such developmental changes are, however, hardly limited to young children; the concepts of adults change over time as they learn more and refine their strategies for recognizing and reasoning about things in day-to-day life. To take one example, I only recently learned how to read safety diamonds (the diamond shaped signs on the back of truck trailers, usually tankers). In learning this, I have acquired new ways of categorizing substances as

¹⁶⁶ For instance, one possibility often endorsed by psychologists is that this shift in categorization strategy is reflective of a shift from the use prototypes/exemplars and similarity-based matching processes to the use of theories and rule-based processes.

specifically flammable, caustic, radioactive, and so on. Similarly, as a teenager I learned new properties of already known chemical substances and came to be able draw the relevant inferences upon encountering these substances, such as that ammonia and bleach when mixed produce chlorine gas and that one must be mindful of this when cleaning. Both my newfound categorization strategies and my new inferential dispositions are grounded in the same structural shift in my concepts, likely an expansion in both my definition for safety diamonds and my folk theory of chemical kinds.

3.2.3 – Differences in Capacities

With regards to (3), differences in fine conceptual structure can arise due to differences in the cognitive, perceptual, and practical capacities had by concept possessors. Hellen Keller will have no concepts which represent properties available only to visual and auditory perception, and will likely possess concepts that represent properties many of us pay comparatively little attention to (such as those available to touch) and likely do not represent in thought or assign low weight to.¹⁶⁷ Someone who is colorblind is unlikely to heavily rely on certain colors of objects when categorizing, such as in cases where they are unable to accurately recognize the presence of those colors or are likely to confuse them with a similar color. And depending on one's genes, one may be unable to represent cilantro as having a taste anything other than gross and soapy.

¹⁶⁷ The exception may be features that she could associate with things via language. She likely, for instance, has a concept RED and RED likely formed part of the structure of concepts of things she had been told were red. Similarly, perceptual and cognitive prosthetics, such as devices that translate detected wavelengths of light into patterns of vibration on the wearers skin could allow identification of, and thus present value in representing, visual properties to those not normally able to detect them. Though in these cases, the fine structures of her concepts would likely differ a good deal from those had by people with normal vision.

With regards to cognitive capacities, my concept of water may represent it as being H₂O while my cat cannot do the same for the water in its bowl. This is not because I can literally see that the water has that chemical structure and my cats' visual faculties are comparatively limited. Rather it is because I am capable of certain kinds of cognition, such as linguistic comprehension and the ability to learn via instruction from chemistry teachers, that my cat is not built to engage in. And limitations in my long-term memory may mean that I have far more impoverished exemplars of the people and events of my past, representing them in lesser detail than those around me with a greater capacity to recall the past.

Finally, some of the properties that can be represented by a thinker's concepts will reflect the practical possibilities that the represented aspects of the world offer to the thinker. A drug-sniffing dog may well represent the drugs it detects as a means for acquiring food or attention. But it will not represent them as offering recreational or monetary opportunities. This is in part because the dog may be unable to understand that the drug affords these possibilities in the first place (though cats clearly learn of the psychoactive effects of catnip without much difficulty). Whatever the reasons, the dog is likely unable to have desires with content, and as I will argue in chapter 7 this will cut off the possibility of the dog representing these aspects of the drug in thought.

3.2.4 – Differences in Context

For (4), consider first differences in context of acquisition. A contemporary person's concept of Aristotle is likely to be far different from that of the Greeks that knew the man personally. Given how we are situated historically relative to Aristotle, about the only way we can possibly come to have any concept of him is via his writings or those who refer to him in either speech or writing of their own. Due to this, there is a very limited number of ways

whereby one can now come to acquire a concept of Aristotle, and this limitation may result in greater uniformity of concepts now than when he still lived. Our concepts can only “hook on” to the man via something like a description or via piggybacking off the semantics of the associated name in some manner.¹⁶⁸ We cannot form a concept of him on the basis of our perception of or direct dealings with the man, for we have none and could not have any.¹⁶⁹

It would not surprise me in the least, for instance, if the vast majority of Aristotle concepts out there now had structures consisting of little more than STUDENT OF PLATO and PHILOSOPHER. The concepts of those Greeks who knew him personally likely represented far different properties. Their firsthand experience means their concepts are more likely to represent the perceptual properties of Aristotle, or at least to assign them a higher importance. And these differences would predictably result in the expected differences in capacities: while Aristotle’s personal friends and family would have no difficulty categorizing the man on the basis of his physical appearance or voice, none of us could do so. They could find him in a busy market, though we could not.

3.2.5 – Differences in Purpose

That the ancient Greeks had concepts with differing fine structures of our own should not be surprising/ Aristotle’s wife had need to recognize his face and voice in a crowd, while we merely need to recognize him based on his philosophical works and when we see or hear his name used. The different ends of his wife relative to most of us means that she needed to

¹⁶⁸ See Millikan (2017), pg. 52-54, for one (somewhat heterodox) suggestion for how such piggybacking may work.

¹⁶⁹ Millikan has proposed that one can perceive objects through words or other signs of them. It is unclear whether this extends to things in the historical past referred to by contemporary tokens of the relevant names, but if so she would be one (and perhaps the only) theorist who would deny that a concept cannot be obtained of Aristotle in this manner.

rely on different information than we do in carrying out categorization and reasoning tasks, and these differences demand that her concepts differ from ours in terms of their fine structures. This point generalizes: differences in the needs or ends of concept possessors will give rise to differences in the fine structures of their concepts. This is most obvious in the case of ideals. Differences in our needs and preferences will give rise to differences in what we represent an ideal member of a kind as being. If I adore rain and coffee, then I may well represent Seattle as an ideal city to live in. But if you abhor the rain and cannot stomach coffee, then Los Angeles or somewhere else bereft of rain and quality coffee may be closer to what you represent as being ideal.

Differences in experiences, epistemic access, capacities, needs, and ends will thus result in differences in the concepts developed on the basis of such experiences and in service of such needs and ends. These differences in concepts will become even worse once we look beyond differently situated human beings and look at creatures far more different from us in terms of their form of life, needs, and capacities. A shark may represent food partly in terms of its electrical properties, but I, lacking electroreception, cannot and will not do so. A gosling during early development may represent its mother solely by a property like LARGE OBJECT, and given how it develops, its normal environment, and that it only need stay nearby its mother for survival, that will work well enough for its purposes. But given my stage of development, my environment, and my needs, I will need a good deal more to work from if I am to reliably recognize and engage with my mother. If other humans are unlikely to share the same finely structured concepts, then chances are far worse that finely structured concepts will be shared by both us and non-human animals or beings with artificial psychologies.

This point should not, however, be taken too far; there are some possible exceptions to the general rule of widespread conceptual variability. For some particulars, the exemplars had by many people may develop in ways that result in them being highly similar. This could be owing to factors such as how those exemplars are acquired, the nature of the things represented, and the epistemic position of the relevant thinkers. To adapt an example from Kripke, it may well be that the only thing the average person believes of Einstein is that he was a physicist who invented the atomic bomb.¹⁷⁰ In this sort of case, we would expect the exemplars for Einstein among most people to be both fairly impoverished (owing to their lack of beliefs about the man) and to be fairly uniform, representing properties like male, physicist, and inventor of the atomic bomb. This is likely owing to facts about how most people, at least in the US, learn about Einstein. Such cases can arise even for exemplars of members of kinds. For instance, there is likely to be a good deal of uniformity between typical thinker's exemplar sets of planets, owing to there simply not being that many to form exemplars of in the first place.¹⁷¹ And other concepts like prototypes may well be widely uniform in some cases as well. For instance, in cases where stereotypes in the traditional sense exist, then one would expect there to be greater uniformity of prototypes between members of that population. It is plausible, to take one example, that the concept NURSE for many in the United States includes features like WOMAN owing simply to cultural factors that shape both who goes into nursing and how nurses are portrayed and talked about in media.

¹⁷⁰ Kripke (1980), pg. 81-87.

¹⁷¹ Though even here there is room for some variation. While most in the developed world likely have exemplars for each of the eight planets in the solar system, some will have richer sets than others due to them having exemplars for notable exoplanets and for fictional planets. A science fiction fan may have exemplars for Arrakis and Gethen, while the less nerdy among us may not. And those bitter over Pluto's demotion may still include it out of spite towards the IAU.

3.3 – Objections

Despite what is, I think, strong intuitive and empirical evidence in favor of pluralism, there is still room to object. In this section I will consider two such objections. First, I will consider the objection that concepts constitute merely a small subset of our knowledge, and that much of the other kinds of information I have argued we need heterogeneous concepts to capture can be captured instead by background, non-conceptual knowledge. The second objection is that instead of a pluralist theory we ought instead adopt a hybrid theory of concept. I will tackle each of these objections in order.

3.3.1 – Background and Conceptual Knowledge

One possibility that may render pluralism unnecessary is that concepts include only some small subset of our knowledge of any given category. If this were the case, then the fact that we draw upon such a wide variety of information when we exercise core cognitive capacities would not in itself tell us what kinds of concepts we possess or relatedly what kind of information those concepts encode. Some of that information, such as that regarding typical properties, may be stored in concepts while the rest is stored in one's background beliefs. How exactly this general idea is developed will depend on what information you think concepts encode and what the precise theoretical role for concepts is, e.g. whether concepts are, as some argue, only used in making quick, unreflective judgments.

During the heyday of the classical theory, a common view was that concepts encode necessary and sufficient conditions while background knowledge concerned merely contingent properties.¹⁷² The knowledge that concepts provided then was privileged, in particular

¹⁷² Compare knowledge that bachelors are unmarried adult men to knowledge that bachelors tend to be messy.

epistemologically, in a way one's other knowledge was not.¹⁷³ On this view, it may be that I rely on the typical properties of a thing when making a categorization judgment, and it may even be proper to call my representation of these properties something like a prototype or a stereotype. But such a prototype is not properly called a concept. Rather it is something like an identification procedure or a mere conception, something which may serve to help to setup and maintain a content-determining relation for that concept, guide that concept's application, or capture the subject's own idiosyncratic understanding of the category rather than some abstract ideal captured by the concept.¹⁷⁴

One could also endorse a distinction between knowledge that is conceptual and that which is background by taking the reverse stance. This would involve holding that concepts encode contingent properties while knowledge of, say, definitions constitutes the background. On this view, concepts are used only in making snap judgments, while background knowledge is drawn upon when making considered judgments, during tasks where greater accuracy is needed or there is more time available for deliberation. A prototype theorist, for instance, may hold that the only kind of concept that there is are prototypes, that we use these solely for categorization (and other relevant tasks) when under strict time constraints, and that we rely on background knowledge only when our prototypes prove to be inadequate to the task. One could take similar stances to the two just mentioned with any combination of the concept kinds

¹⁷³ It is for this reason that philosophers engaged in conceptual, rather than mere belief, analysis. Conceptual knowledge was thought to have guaranteed accuracy and robustness in a way that the rest of one's beliefs lacked.

¹⁷⁴ This last take only really makes sense if one thinks of concepts as some philosophers do, as things more akin to abstract Fregean senses than mental representations which underly much of our cognition.

discussed thus far, e.g. concepts being exemplars while theories constitute our background knowledge, or with either concepts or background being heterogeneous in some manner.

Regardless of what way one carves up the cognitive domain, one requirement universal to all carvings in this general spirit will be that there must be a robust distinction to be drawn between the information encoded in concepts (whichever ones one endorses) and that information encoded in the purported background knowledge, i.e. a distinction between what is often fittingly called conceptual and background knowledge. In what follows, I will try to make clear what this distinction amounts to and argue against its explanatory value.

Before considering the plausibility of such a distinction, it is important to first make clear that the background/conceptual knowledge distinction, while similar to the analytic/synthetic distinction, need not be identical to it.¹⁷⁵ For one, whether there is a background/conceptual knowledge distinction is an empirical matter, while few I suspect would take the issue of whether there is an analytic/synthetic distinction to be the same.¹⁷⁶ There is nothing metaphysically impossible about a cognitive system that draws upon prototypes in every context and for every purpose, and reserves, say, bare sentences in a language of thought for cases where the prototypes have proven to be inadequate or unreliable for one reason or another. There is nothing impossible, then, in a cognitive system that is bifurcated into two subsystems or representational mediums that differ greatly in their functions and the tasks for which they are utilized, regardless of whether our cognitive systems

¹⁷⁵ Whether it the two distinctions are one in the same depends at least in part on how you understand each. My point in is not that they in fact are distinct. Rather, it is that on certain understandings of each distinction the two can come apart, and thus the identity of the two distinctions should not be assumed.

¹⁷⁶ One exception is Rey (2014), pg. 93-94.

are actually like this. Conversely, there is nothing impossible in a cognitive system that dispenses with concepts as understood here and utilizes nothing but sentences in a language of thought as its cognitive representations. Both are live possibilities, at least before the empirical evidence begins to be considered and the space of theoretical possibilities begins to contract. Few philosophers would, I take it, hold that whether an analytic/synthetic distinction exists turns on facts such as these about human cognitive architecture and processes.

Second, it may be that the analytic/synthetic distinction, suitably understood, is one relevant to the meanings of words in a natural language, without these word meanings being related in any clean or systematic way to the cognitive content of concepts. I will illustrate this point through the example of a specific theory, but I believe other theories of language could take a similar line. Consider the teleosemantic theory of language proposed by Millikan. She holds that linguistic forms have proper functions. For instance, the proper function of the indicative mood is to produce belief in the hearer, while the proper function of the imperative mood is to induce the hearer to act in ways indicated by the imperative.¹⁷⁷ Speakers continue to use imperatives because it brings about effects, affected by the hearer, that the speaker desires; hearers respond to imperatives because it brings them rewards of various kinds (or at least avoids possible negative consequences). It is these effects of the indicative and imperative moods which account for the continued existence of these forms within linguistic communities. Were imperatives never followed or if indicatives never induced belief, their speakers would

¹⁷⁷ Millikan (1984), ch. 3.

cease using them and they would soon die out.¹⁷⁸ A radically individualistic society (say one of Randian men of industry) may have no use or desire for the imperative mood.

Similarly, words have functions of their own, what Millikan terms “stabilizing functions.”¹⁷⁹ The stabilizing function of a word is what that word has done historically that accounts for its continued use by speakers. This continued use in turn requires that hearers respond in the relevant ways. The word “fire,” to take one example, will only continue to be used if speakers can achieve their purposes via the use of the word (in part) to induce beliefs in hearers (to a child, that fire is dangerous) and actions (to warn a person to flee from a burning building). And speakers can only achieve these purposes if hearers respond appropriately (children coming to believe that fire is dangerous; those warned actually fleeing the building). In order for words to perform these stabilizing functions, they must also function to allow their possessors to identify the world-aspects to which they refer. If you command that I close the window on a chilly day, that command can only achieve its function of producing my compliance if I do in fact comply. And my actually complying requires, among other things, that I be able to identify the window that you want me to close, and thus am able to identify what the word “window” in your command is referring to.

Importantly, these stabilizing functions of words can remain the same despite variations in how the word is understood by its users (provided the variation is not too severe). The word “fire” could, for instance, be useful to its users in giving and heeding warnings (among many

¹⁷⁸ Importantly, this does not require that they achieve these functions all or even most of the time. All that is required is that they achieve these effects often enough that they continue to prove useful for a critical mass of language users. And it may be that some of us, say those in position of low social standing, are not positioned to ever use forms such as the imperative in ways such that it performs its proper function.

¹⁷⁹ Millikan (2005).

other possible uses), even if what exactly fire is understood to be differs between the people using that word. If Aristotle were to know English, he could understand a call to flee a burning building even if he understands fire as one of the four basic elements and the one giving the warning understands fire in terms of oxidation and the chemical process of combustion.

While Millikan denies that there are meanings in the sense relevant to analyticity, this does not seem to be something that teleosemantic theories of language more generally must be committed to. A word fulfilling its stabilizing function requires that a critical mass of its users be proficient in those usages which satisfy speaker and hearer purposes. And it may be that such proficiency in these usages requires associating with the word a common set of concepts or concept elements which may be considered its meaning in some sense, e.g. it may be that “bachelor” could not perform its stabilizing function and stay in circulation unless a critical mass of users associated UNMARRIED, ADULT, and MAN with the word. And this common association could be what accounts for the seeming fact that competent users of the word “bachelor” can know upon mere reflection that bachelors must be unmarried.

On such a view it may make sense to draw a distinction between those concepts/concept elements that are relevant to the performance of such a word’s stabilizing function, and those that are relevant to its more idiosyncratic uses.¹⁸⁰ And this could potentially supply grounds for considering the former concepts/concept elements part of the meaning of and thus analytic to the word, and all other concepts/concept elements concerning bachelors synthetic. But this

¹⁸⁰ Compare the properties of a screwdriver that allow it to drive screws (those relevant to its proper function) and those that allow it to perform more idiosyncratic things, such as prying open paint cans. Screwdrivers may only stay in production so long as there are screws to screw, even if screwdrivers are put by their users to other productive uses on occasion. People open paint cans with screwdrivers because they have previously bought screwdrivers; they do not buy them for that purpose.

on its own need not imply anything about the architecture of the minds of those in that linguistic community who make use of that word. It may even be that what counts as the meaning of the word cuts across multiple conceptual kinds, with part of the word's meaning being derived from information in an associated prototype and another part from an associated theory, e.g. UNMARRIED from a theory and MAN from a prototype.¹⁸¹ In other words, that a word has a meaning derived from the cognitive content of concepts implies nothing on its own about the structure of those concepts or the broader cognitive architecture that encompasses them. It does not even imply that the word's meaning is derived from a single concept as opposed to many.

Finally, there is nothing in the conceptual/background knowledge distinction that implies there is anything special about conceptual knowledge beyond its being used by default. Of particular relevance here is that there is no requirement that conceptual knowledge be true (remember the idiosyncratic use of "knowledge" by many of those engaged in these debates). It may be that it is part of the ordinary person's concept of WHALE that whales are fish. This may be "knowledge" retrieved by default and relied upon across all contexts. And thus it could be that (1) one can demarcate conceptual from background knowledge, and that (2)

¹⁸¹ This possibility is easier to accommodate on theories like those in cognitive linguistics that deny the conceptual/background knowledge distinction. On the theories in that tradition, it is thought that linguistic comprehension involves bringing to bear potentially any part of one's knowledge of the world, with which part being determined by things like the context and ends of the language users involved. See Evans, Green (2006), ch. 5 & 7, for an introduction of this approach to linguistic meaning. For the record, I do not endorse this idea, and in fact think it is likely false. The point is merely that there can be concepts associated with a word in a privileged way such that it may make sense to say they constitute its meaning in some sense while simultaneously failing to be associated with each other in the mind in ways that matter for the purposes of most psychological explanations.

“knowledge” of that first kind can be false. But again, I take it that few philosophers would be willing to accept that there are positive analytic claims that are false in this way.¹⁸²

Setting these issues aside, here is the main question: Is there a conceptual/background knowledge distinction? Is there reason to believe that there is a distinction between that information that is encoded in one’s concepts and that which is encoded in one’s background knowledge? The cognitive psychologist Barbara Malt gives two reasons for saying “no” on both points.¹⁸³ The first of these is that there is no sharp cutoff point to be found where subject’s shift from using information that could be termed conceptual to using a distinct body of information that could be considered one’s background knowledge. If there were a distinction between representations encoding each type of information, there would need to be some way that the distinction manifests in thought and behavior. If it did not, then the distinction would be explanatorily idle and positing it unmotivated. And the main difference psychologists have expected to find is that the information in concepts is that information that is relied upon by default, reliably drawn upon spontaneously and in an acontextual manner to guide behavior.¹⁸⁴

If there were a background/conceptual knowledge distinction then, we would expect something like the following picture. Subjects first retrieve some specific body of knowledge which is always deployed when dealing with whatever world-aspect they are concerned with. They then attempt to make their categorization, draw their inference, or whatever else they are trying to do. If this knowledge proves adequate to the task, no additional knowledge is drawn upon. However, if the knowledge initially deployed instead proves inadequate in some way,

¹⁸² One exception is Rey (2014), who argues that this is the case.

¹⁸³ Malt (2010).

¹⁸⁴ This is a point I will return to later in ch. 5 when considering Machery’s (2009) definition of concepts and what I take to be issues with his understanding of the function of concepts.

then additional knowledge is pulled from long-term memory to serve as a supplement. On this view, that initial body of knowledge retrieved is what is supposed to be conceptual, while the backup knowledge held in reserve and deployed if needed is supposed to be background.

Malt argues, however, that there is no clean line to be drawn between a supposed conceptual stage of cognitive processing and a stage that draws upon something else, such as background knowledge. Subjects engaged in a task appear to be constantly accessing more information the longer the task goes on, rather than there being two discrete stages characterized by different kinds of representations and processes. Further, there does not appear to be any explanatory value to be had in drawing a line between the two kinds of knowledge that would be drawn upon by distinct cognitive processes:

Consider the knowledge that is retrieved when hearing a word such as *chair*. What knowledge comes to mind depends on how long a person has to think about it. In the first 20 seconds, what comes to mind may be that chairs have 2 arms, 4 legs, and are for sitting. In the next 20 seconds, more information will come to mind: maybe that chairs can be made of wood or metal, can be upholstered, can be stuffed, and can be used at a table or in a living room. And with further thought, what comes to mind may also include that they can be made of plastic or aluminum, can have no arms, can be used at the beach, can have casters instead of feet, can have rockers, and so on. To define a default, there would need to be a cutoff for what counts as the default or spontaneously retrieved information. It is not clear if there is a non-arbitrary way to make that decision.¹⁸⁵

¹⁸⁵ Malt (2010).

Secondly, Malt notes that what information is first retrieved in the performance of tasks like categorization is a highly contextual matter.¹⁸⁶ Further, retrieval from memory more generally appears to be highly context-dependent as well.¹⁸⁷ To take one example it may be thought that there is a single body of knowledge that can be thought to characterize ART as a concept, one deployed in all contexts and supplemented only as necessary. Works of art, however, vary drastically from one another. They can differ depending on the artist's own personal style, the art movement the work belongs to, the time and place it originated in, and so on. What knowledge I have of art that is retrieved when looking upon or thinking of the Mona Lisa is likely to be quite different from that retrieved when thinking about Marcel Duchamp's Fountain. The former may conjure to mind my ideas about the beauty of art and how it engages one's aesthetic senses. It may bring to mind prototypes such as that of a brightly lit, clean museum, filled with tourists and young students on class tours. It may bring to mind exemplars of other famous works of art, like Michelangelo's Pieta or Botticelli's The Birth of Venus.

By contrast, with Duchamp's fountain I will likely draw upon knowledge of quite different matters. The work is ugly and off-putting by design, a urinal placed in a setting unfit for one. By design it is meant to force consideration of one's definitions of art, or one's theory of the function and role of art within society or the artistic community. It may cause the retrieval of exemplars, same as the Mona Lisa. But the exemplars retrieved are more likely to be Dadaist works than classical or impressionist ones. In fact, were one to retrieve the same

¹⁸⁶ Barsalou (1982), Roth & Shoben (1983).

¹⁸⁷ Godden & Baddeley (1975).

knowledge when confronted with Fountain as one did when encountering the Mona Lesa, there is a high likelihood that one simply did not “get it.”



Figure 3.1¹⁸⁸

Whether there is a conceptual/background knowledge distinction, is however, still ultimately an empirical question. As is the question of whether the processes supporting core cognitive capacities can be split into two discrete stages, or whether knowledge retrieval is as contextually dependent as suggested here. As I said earlier, there is no impossibility in either proposal for any of these questions. However, what I suspect is the case, as much evidence currently seems to indicate, is that there is no privileged knowledge that is deployed by default when engaged in tasks like categorization and other forms of reasoning, this knowledge only supplemented once its proven inadequate to the task. Actual cognition in the creatures that we are familiar with appears to be less regimented and far messier than such a picture suggests.

¹⁸⁸ Original photo by Alfred Stieglitz (1917).

Of course, there is no impossibility in a creature that divides its knowledge of the world in two, drawing upon one set of knowledge in a privileged manner. And such a creature may relegate one kind of knowledge to rapid and spontaneous thought, and the other to careful, considered thought. Evolution could surely build such a creature if its environment and needs demanded it. And we could presumably build an artificial mind that operated in this manner if our plans dictated it. Such creatures are possible. The point here is merely that we do not appear to be such creatures.¹⁸⁹

3.3.2 – Hybrid Theories of Concepts

The hybrid theory supplies another objection to pluralism. The hybrid theorist can agree with the concept pluralist that we utilize all the sorts of information encoded by each of the major kinds of concept when we exercise our core cognitive capacities. They can further agree that there is no meaningful distinction to be drawn between conceptual and background knowledge. However, rather than saying that this medley of information is all stored in a multitude of distinct concepts, they can instead say that it is stored in distinct but integrated *parts* of a single *unified* concept. Given this possibility, I will say a few things in defense of a pluralist approach to concepts over a pure hybrid one.

One difficulty the hybrid view faces is specifying what kind of evidence would indicate the presence of hybrid concepts as opposed to distinct concepts corresponding to each of the hybrid's supposed parts. For instance, suppose that the hybrid concept posited by some theory is composed of a definitional core and a prototype at its periphery. Thus it has one part that represents the necessary and sufficient conditions for category membership in the represented

¹⁸⁹ Nor, I suspect, are many (if not most) of the non-human animals that we may consider.

kind, and another part that represents the typical or cue-valid properties of the members of that kind. But how can we know that such a hybrid concept is being retrieved and made use of, as opposed to both a definition and a prototype being retrieved independently of one another?

One proposal is given by Machery, who argues that hybrid concepts are distinguished from collections of distinct concepts via the information they encode being both *linked* and *coordinated*.¹⁹⁰ To say that the bodies of knowledge are *linked* is to say that when one body of knowledge is retrieved and made use of by one's cognitive processes, the other body of knowledge is thereby made available for immediate use by other processes. To say that the bodies of knowledge are *coordinated* is to say that they do not give rise to inconsistent judgments, e.g. one body of knowledge giving rise to the judgment that some world-aspect W belongs to a kind K, while another body of knowledge gives rise to the judgment that W is not a member of K.

Whether or not the hybrid theory is ultimately true of all concepts across all thinking things is again an empirical matter. There is no impossibility in either (1) the pluralist view that there are distinct conceptual representations that are deployed in ways sensitive to task and context, and (2) the hybrid view that there is a single representation with multiple parts, with each *part* deployed in ways sensitive to task and context.¹⁹¹

While this is ultimately not a question that philosophy is well-suited to answer on its own, there are a couple of considerations that I believe ought to incline one towards a pluralist over a hybrid position. For one, hybrid concepts are, by definition, more complex representations than prototypes, definitions, exemplars, and such. They would have to be,

¹⁹⁰ Machery (2009), pg. 63-66.

¹⁹¹ It is also possible, of course, that neither pluralism nor the hybrid theory is true.

given that they are compounds of these simpler concepts. The processes that make use of them would have to be more complex as well, with distinct parts of the process or distinct processes entirely corresponding to each part, e.g. a process of checking definitional fit for the conceptual core and a process of similarity-matching for the prototype periphery. This greater complexity in both conceptual structure and processes carries with it clear benefits. One such benefit is that it prevents making overly simplistic categorizations solely on the basis of superficial properties, which is valuable as such categorizations can often go wrong in many contexts.¹⁹² It also, however, comes with clear tradeoffs, such as imposing a higher cognitive load and requiring processes that demand more time to execute.

These tradeoffs may matter in two ways. First, there will be certain tasks and contexts where a simpler concept is perfectly adequate and more economical to utilize. Given this, it is reasonable to expect that we (or at least some thinking things) would develop these simpler concepts and be able to utilize them in isolation for when such tasks and contexts arise. Second is that more complex, hybrid concepts are harder to acquire and impose stricter and more difficult to satisfy conditions for possession. For some thinkers, such as more cognitively simple animals or very young human children, these conditions may be impossible to meet. There is little to no evidence, for instance, that young children utilize definitions, and as discussed previously they do not appear to start developing mental theories until around age five. Prior to this the child must be thinking about the world and using something to categorize and draw inferences about its contents. Plausibly, what the child is using during these early stages, where they are sensitive only to the superficial, perceptually available properties of

¹⁹² How likely they are to go wrong under *normal* contexts, the sorts of contexts concepts like prototypes tend to actually be used in by their possessors, is an issue that will be discussed in more detail in ch. 7.

things, are prototypes and exemplars. And the evidence does seem to show that children are relying primarily on similarity of superficial properties at these young ages.¹⁹³

Second, and more importantly, is that there is no incompatibility between the hybrid and the pluralist positions. A pluralist can happily admit that some of the many kinds of concepts which are or may come to be possessed are hybrid in their coarse structure. Perhaps young children do make extensive use of prototypes and exemplars, with these being developed into more complex hybrids of prototype, exemplar, and definition as they grow older. It may even be that as adults we utilize both a bare prototype and hybrids including that prototype depending on task and context. This is, again, ultimately an empirical question, and one that seems to be currently quite undecided. Nor is it seemingly clear how to determine empirically which is the case, which is why there is a need for proposals such as Machery's on what makes bodies of knowledge part of a hybrid deployed singularly in service of a task as opposed to multiple bodies of knowledge deployed jointly. What is clear is that we cannot determine a priori which of these concept types we and other creatures we know in fact make use of. But it may be that we can know that these representations are all *possible*, that there is nothing about what makes something a concept that rules them out.

Given that pluralism can accommodate any representational kind that hybrid theorists may offer, along with the daunting task of showing that hybrid concepts can accommodate all that has been expected of concepts generally, I see little reason to prefer a pure hybrid approach over a pluralist one. The pluralist can happily accept the representations and processes posited

¹⁹³ However, see Keil (1989), in particular ch. 10-12, for arguments that very young children do in fact rely on theories for at least some domains. Though for those domains that adequate theories have not yet been developed for at those ages, my point would still stand.

by the hybrid theorist into their ontology along with all the explanations provided by them. The hybrid theorist cannot do the same with those representations posited by the pluralist. All the pluralist demands in addition is that the hybrid theory not be both the beginning and the end of the story when it comes to concepts.

4 – Concept Individuation: Atomism Reconsidered

Much of the discussion up to this point has concerned debates over the structure of concepts and the roles that such structure plays in explanations of various kinds. From all this, it may seem that positing such structure and its accompanying cognitive processes is necessary for concepts to fulfill their explanatory role. And from this it may seem then that such structure and processes are inarguably an essential part of any plausible theory of concepts. Perhaps somewhat surprisingly, this is not the case, or at least has not seemed to be so to some philosophers. There is in fact substantial debate between philosophers over the roles such structure plays and even doubts about its very existence.

In what follows I will consider the two sides of this debate over the role and existence of conceptual structure: conceptual molecularism and conceptual atomism. Molecularism roughly holds that concepts are structured entities, however this claim is understood.¹⁹⁴ Consider Fodor's test for whether you may be a molecularist:

If you hold that it's not possible to have the concept RED unless you have the concept COLOR, then your theory of concept possession is to that extent an-atomistic [molecularist].¹⁹⁵

It has long been and remains the dominant position in both philosophy and psychology. Atomism, by contrast, roughly holds that concepts are unstructured, and is largely limited to philosophy in terms of support. Given an understanding of both what concepts are and what it means for them to be structured, the difference between the two positions at first seems stark and obvious. Molecularists simply believe that concepts are structured, either by having literal

¹⁹⁴ See Fodor (2004a), pg. 35, for an understanding of it in terms of individuating inferential connections.

¹⁹⁵ Fodor (2003, pg. 147).

parts or by standing in inferential/functional/causal relations to other concepts, while atomists simply deny this claim. Atomism seems then to be a straightforward negative theory, defined solely in terms of its denial of the molecularist's defining position.

Much of my goal here will be to make the case that this debate, and in particular the theory of atomism, is less clear than it is often characterized. If atomism were nothing more than the view that concepts are unstructured, then atomism would be a complete non-starter. As has long been argued, unstructured concepts cannot do anything, least of all encode information about one's environment and guide complex action.¹⁹⁶ The structure of a concept is what encodes the information in the concept, and it is a concept's structure that cognitive processes exploit and make use of to guide the exercise of core cognitive capacities. Further, atomism so understood would require the denial of the current leading theories of concepts in cognitive science, all of which posit structured conceptual representations and cannot eliminate such structure without rendering those theories explanatorily impotent. This alone would make such a position unpalatable to any naturalistically inclined philosopher. And naturalistic inclinations, as a general rule, are a mark of a philosopher being more, rather than less, inclined towards atomism as a theory of concepts.¹⁹⁷

For these reasons and others, I will argue here that the debate, despite first appearances, is not over the existence of conceptual structure at all. Rather, the debate is over both (1) what kinds of structure concepts possess, and more centrally (2) what status those structures have. More particularly, I will argue that the debate is really concerned with whether the fine-grained

¹⁹⁶Laurence & Margolis (1999), pg. 64-65; Prinz (2002), pg. 99-100; Prinz & Clark (2004); and Daley (2010).

¹⁹⁷ This is understandable given that one mark of naturalism is a disposition (albeit limited) to defer to successful scientific theories, of which the myriad theories of concepts in psychology are, I take it, clear examples.

representational structures that concepts have, i.e. the specific contentful representations had by them as parts or to which they stand in explanatorily privileged inferential/causal/functional relations, are individuating of them.¹⁹⁸ To put it another way, the debate is not over whether concepts are structured so much as whether concepts must be structured in particular, fine-grained ways in order to fall under explanatorily relevant concept types.¹⁹⁹

This chapter is concerned primarily with making sense of what both sides of this debate are (or perhaps better *should* be claiming), rather than arguing that one of the sides is correct. While my own sympathies lie closer to atomism, in chapter 5 I will argue that neither molecularism nor atomism ultimately succeeds. This is not because of anything directly tied to the specifics of the debate covered in this chapter. Rather, it is due to a deeper instability of the notion of concept that both positions are founded upon. And as I will argue, a more radical position than either molecularism or atomism will ultimately be needed.

4.1 – What is Molecularism?

As mentioned above, molecularism in the simplest of terms is the view that concepts are structured entities. Though molecularist concepts are usually understood as structured mental representations by many philosophers in the literature and nearly all psychologists, the molecularist need not be committed to any specific position on the ontological category that concepts belong to.²⁰⁰ Consistent with my earlier assumptions, however, I will continue to

¹⁹⁸ For the sake of simplicity, I will refer merely to inferential relations from here on out.

¹⁹⁹ Note that one can hold that concepts are structured without holding that their structure plays any role in their individuation. Despite much discussion, this distinction is still sometimes ignored in the literature. See the discussion of atomism from Laurence and Margolis (1999), pg. 71-75, for arguments regarding the importance of this distinction.

²⁰⁰ For instance, Peacocke is a molecularist who denies that concepts are mental representations, instead arguing that they are mere abstracta of some sort. Similarly, Millikan's (2000) early view on concepts can be plausibly construed as molecularist in this simple sense, though she holds that concepts are abilities rather than representations.

understand concepts as mental representations in what follows. While this definition of molecularism is simple, easily stated, and common in the literature, I will argue in what follows that it is inadequate and fails to properly distinguish molecularism from competing proposals.

Prior to this, however, it is important to first consider why molecularism has long been appealing to both psychologists and philosophers. It is worth lingering on these issues for a time. The reasons for molecularism's popularity are numerous and initially compelling. Understanding the longstanding appeal of molecularism will serve to guard against a too quick dismissal of it upon first consideration of its problems, while also rendering understandable (if not justifiable) the equally quick dismissal of its atomistic alternative by molecularism's proponents. For it is the seeming inability of atomist concepts to do the work molecularist concepts accomplish easily that forms atomism's single greatest stumbling block.

4.2 – The Appeal of Molecularism in Psychology

As demonstrated and discussed in the previous chapters, conceptual structure seems to be indispensable for the purposes of psychological explanation. This is, in large part, why molecularism is not just the dominant view of concepts but also the default one. As this is well-tread ground (including within this dissertation throughout chapters 1 and 2), my comments here will be brief and primarily serve as review.

In psychology the primary focus for studies of concepts has been the explanation of core cognitive capacities. These include, but are not exhausted by, capacities such as categorization, various kinds of reasoning, and linguistic comprehension. Explanations of these capacities has traditionally depended upon concepts having structure of some kind or another. Taking a toy example, an explanation of a subject's recognition of an apple presented to them in perception will depend upon their APPLE concept being related to concepts and/or

perceptual representations of the properties on the basis of which the categorization judgment is made. This categorization proceeds in part from the prior recognition of the apple's redness through the related concept RED or through a perceptual representation of redness. Were APPLE isolated from other concepts and representations of this sort, there would be nothing to guide the application of it. Without reference to relations of this sort, i.e. without reference to the concept's structure, explanation of the subject's recognition is seemingly impossible. Thus, molecularism's popularity in psychology largely flows from the accompanying promise (and indeed reality) of successful explanations of the phenomena of most interest to psychologists.

By contrast, those theoretical roles which structured concepts have been less suited for have received correspondingly less attention by psychologists. That structured concepts seemingly struggle to account for the publicity and stability of concepts over time are problems comparatively ignored.²⁰¹ Relatedly, psychologists have seemed largely uninterested in externalist theories of mental content, leaving only internalist theories that (when paired with widespread conceptual variation, ignorance, and error) would lead to radically different extensions from thinker to thinker. In general, the semantics of concepts is rarely, if ever, considered by most psychologists; that concepts represent what they do, usually whatever the word expressing them refers to, is simply assumed without further explanation or inquiry.²⁰² This has the effect of masking the more serious problems molecularism faces on that front, thus undercutting one of the main motivations for considering alternatives like atomism.

²⁰¹ There are some cases where issues of publicity are considered, such as Smith, Medin, & Rips (1984). However, in these cases the focus is exclusively on publicity for cognitive content. Publicity for referential content (and issues concerning reference generally) are rarely discussed, or even sometimes have their relevance outright denied (see Hampton 1999).

²⁰² Hampton (1999), Malt et al. (2011).

4.3 – The Appeal of Molecularism in Philosophy

In the earlier discussion of the classical theory of concepts two reasons have already been mentioned for molecularism's popularity in philosophy.²⁰³ First, structured concepts may admit of analysis and thus serve as a possible ground for a priori knowledge and make possible a philosophical methodology founded on conceptual analysis.²⁰⁴ Assuming that concepts are structured (and given several other assumptions, ones quite controversial and not often enough acknowledged), one can acquire knowledge of the world through an analysis of the structure of one's concepts. Atomistic concepts, lacking structure and thus lacking anything admitting of analysis, do not provide such potentially lucrative philosophical possibilities.

Second, structured concepts may provide the basis for reference determination on internalist theories. If a concept's referent is fixed by something internal to the thinker, a natural suggestion is that it is the concept's structure that does the fixing. Thus on this view my concept APPLE refers to apples because it is only apples that satisfy the properties that form APPLE's structure, i.e. RED, FRUIT, GROWS ON TREES, etc. The way of thinking precedes what is thought of in that way. An atomistic concept, seemingly embodying no way of thinking of what it represents and acting as something akin to a mere mental tag for the represented object,

²⁰³ The classical theory does a better job of satisfying these philosophical desiderata than competing views like the prototype and theory-theory do. What is relevant for my purposes here is that it seems that having structure of some kind is a necessary condition a theory of concepts satisfying these desiderata, even if it is true that different kinds of structure make this satisfaction more or less likely.

²⁰⁴ Peacocke (1996) holds this possibility as one of the major advantages of his own molecularist theory of concepts.

seemingly lacks the resources needed to give an internalist story for how concepts come to represent what it is they represent.²⁰⁵

There are of course reasons beyond these two for molecularism's popularity in philosophy. In what follows I will consider two of these reasons: that (1) structured concepts may allow a response to classic Frege cases for thought, and (2) that structured concepts can explain various normative demands that concept possessors seemingly must meet.

4.3.1 – Molecularism and Frege Cases

A third reason many have thought concepts need to be structured is that it has been thought necessary to explain both classic Frege cases in language as well as their psychological analogues. If concepts were unstructured, then the only plausible thing that would make a concept the kind of concept that it is would be what it refers to, i.e. concepts would need to be individuated in a purely referential manner. But if this were the case, then there would be no difference between the concepts CLARK KENT and SUPERMAN, and thus no difference between the representations CLARK KENT IS SUPERMAN and CLARK KENT IS CLARK KENT which contain those concepts as constituents and which support the relevant beliefs that must differ in their content to explain the former's cognitive significance and the latter's lack of the same. This would be problematic, as there are clear behavioral differences between possessors of the corresponding beliefs: Lois Lane presumably believes that Clark Kent is Clark Kent, but her behavior would be substantially altered if she came to believe that Clark Kent is Superman.

Another issue (and yet another motivation for molecularism) is that if concepts were individuated in a solely referential manner, then there would be a difficulty in capturing the

²⁰⁵ This is assuming, of course, that mental content externalism is false. It is likely not a coincidence that all the major conceptual atomists of which I am aware are externalists.

content of representations which include empty concepts, such as UNICORNS HAVE HORNS. The referent of UNICORN would be the same as that of any other empty concept, such as PHLOGISTON or SHERLOCK HOLMES, namely nothing. Thus, there would be no means of distinguishing between such empty concepts, nor any way of distinguishing otherwise identical beliefs, desires, and the like, containing them as constituents.²⁰⁶

A common way of dealing with problems such as these is to hold that concepts have a layer or aspect of content that goes beyond mere reference. Such content is often termed *cognitive content* in the literature. It is cognitive content which captures the particular perspective of the thinker, the way they think about whatever is represented by the concept. This particular perspective or way of thinking is, in turn, often thought to be given by or identified with the structural features of the relevant concept or concepts. For instance, the reason “Clark Kent is Superman” is informative for Lois Lane, while “Superman is Superman” is not, is that her concepts CLARK KENT and SUPERMAN differ in structure despite agreeing in reference. Perhaps SUPERMAN has as part of its structure the concept FLYING while CLARK KENT does not, while CLARK KENT has as part of its structure MILD-MANNERED while SUPERMAN does not. Thus, the beliefs Lois forms upon hearing “Clark Kent is Superman” and “Superman is

²⁰⁶ Issues of this sort intersect a great deal with psychological explanations of core cognitive abilities. For instance, molecularist’s can give a story for why Lois Lane can reliably categorize Superman as such on the basis of properties like *having a cape* and *flying* but not reliably categorize Clark Kent on the same basis. Relatedly it provides an explanation for why she will infer the possession of the property *flying* but not *mild-mannered* following the application of SUPERMAN. Similarly, psychological issues concerning conceptual stability across development intersect greatly with semantic issues in philosophy. Thus, there is at least some artificiality in the division of molecularism’s appeal between psychological and philosophical considerations; issues of concern to psychologists are often indirectly and unknowingly concerns of philosophers and vice versa.

Superman,” those supported by the representations CLARK KENT IS SUPERMAN and SUPERMAN IS SUPERMAN, differ in their structure despite being about the same thing.

This difference in structure between the concepts and the recognition of the identity of their referents (expressed in the belief supported by CLARK KENT IS SUPERMAN) allows the extension of the perspective or way of thinking about Superman to who Lois takes to be Clark Kent, and vice versa. Upon recognizing Clark Kent, for instance, Lois can immediately infer that he can fly, not because this is part of the structure of her concept CLARK KENT, but because it is part of the concept SUPERMAN which is recognized as referentially identical with CLARK KENT.²⁰⁷ Thus, it is no surprise that Lois Lane behaves differently when she believes both CLARK KENT IS SUPERMAN and SUPERMAN IS SUPERMAN, rather than the latter alone. The behavioral difference is traceable to a difference in structure between the concepts and how those structural differences factor into the cognitive processing underlying her core cognitive capacities.

This concern with explaining differences in cognitive significance ties into the earlier concern to explain core cognitive capacities, thus tying together core concerns of both psychologists and philosophers. This is because two concepts can seemingly have the same referent while differing wildly in terms of the cognitive capacities they confer on their possessors. It is because of the differing structure of her concepts CLARK KENT and SUPERMAN that, upon seeing some person flying in the sky, Lois is disposed to infer the presence of Superman (as well as birds and planes), but not the presence of Clark Kent. Plausibly Lois’

²⁰⁷ This does assume that the concepts SUPERMAN and CLARK KENT remain distinct after her realization of the identity of the man referred to by each. What likely happens eventually after such a realization is made is that the two concepts are merged in order to reduce the redundancy present in cognitive systems and the issues which such redundancies may give rise to.

exemplar of Superman consists of featural representations such as RED AND BLUE SUIT, FLYING, and SHOOTS LASERS FROM EYES, none of which are part of her exemplar of Clark Kent. It is not just accounting for differences in cognitive significance then that structure is needed for. Structure is seemingly necessary if we are to account for real differences in the behavior of thinkers and where explanations that appeal to the referential properties of concepts alone prove insufficient. These behavioral differences appear to be the very ones that the models of cognitive capacities developed by psychologists need to be sensitive to.

4.3.2 – Molecularism and Normative Demands on Concept Possession

Finally, conceptual structure is often appealed to in explaining certain normative demands that potential concept possessors must first meet. It is this supposed explanandum that I will dwell on the longest, as it is this which I think has the most far-reaching effects on how many philosophers think of concepts, even if (as I suspect) it is so ingrained in our thinking about concepts that it often goes unnoticed.²⁰⁸ And if such normative demands are real and concept possessors must meet them, then it is the explanandum which will most greatly constrain the shape that a successful theory of concepts can take.

As noted by Mark Wilson, our normal attributions of concepts are sensitive to considerations of what he terms “conceptual preparedness.”²⁰⁹ Whether or not someone is held

²⁰⁸ The comparable issues in the philosophy of language identified by Kripke, Putnam, Burge, and others are seemingly much better known. As I will argue later, many of the externalist arguments given for language apply with little adjustment to thought as well, with many of the same consequences, even if this goes comparatively unrecognized.

²⁰⁹ See Wilson (2006), pg. 23: "In my estimation, a chief service rendered by words like ‘concept’ and ‘attribute’ is that they provide a vocabulary that allows us to monitor and correct our usage as we slowly advance them towards increasingly demanding standards of adequate performance. To fulfill this function sensibly, our talk of ‘concept grasp’ et al. must display considerable sensitivity to the maturational level of the speakers we attempt to evaluate."

to possess or grasp a concept is often a matter of whether or not they can do what is expected of typical concept possessors.²¹⁰ Someone who reliably judges red objects to be blue or circular objects to be square is taken to not plausibly grasp the concepts RED, BLUE, CIRCULAR, or SQUARE on this way of thinking. A similar tying of concept possession to the possession of various capacities is a common theme in the concepts literature, often recognitional abilities of some kind.²¹¹ Indeed, it is so common that Fodor has long argued that this tying together of concepts with practical capacities, what he terms “concept pragmatism,” is something akin to the “original sin” of twentieth-century analytic philosophy.²¹² While this is likely hyperbole on Fodor’s part, there does seem to be a real tendency to view those thinkers who are in some manner inept at reasoning about or recognizing aspects of their world as lacking the concepts corresponding to these aspects. Indeed, it is the lack of the concept (or an inadequate grasp of it, whatever this may mean) that is often appealed to in order to explain such ineptness in categorizing and inferring.²¹³

Conceptual structure explains such normative demands by making possession of a concept require the very structure that confers on one the relevant capacities. Provided that

²¹⁰ What exactly “grasp” means in this context is unclear. The most natural suggestion is that to grasp a concept is simply to possess it. However, grasping comes in degrees while possessing does not; it makes sense to say that someone partially grasps a concept (or, at least, people do talk in this manner), but it is not clear what it means for someone to partially possess a concept. This problem is exacerbated if we make the common assumption that concepts are mental representations realized in the brain, as it makes even less sense to say that someone only partially has a particular mental representation and thus only partially is in a given neural state. Given the obscurity surrounding talk of grasping, I eschew it entirely and use only the notion of possession.

²¹¹ See Prinz and Clark (2004), Millikan (2000), and Machery (2009, 2010).

²¹² Fodor (2004a).

²¹³ It is almost certainly the case that this ineptness is due to *something* conceptual. What I doubt here is whether it is the lack of the concept entirely, rather than one’s concept and its encoded information being inadequate or insufficient to the task, which explains this.

concepts are essentially structured, concept possession can be tied, if not definitionally then near enough, to competency in engaging with the aspect of the world represented by that concept. If possessing SQUARE requires that I find natural and compelling the inference from *S is a four-sided closed figure with sides of equal length* to *S is a square*, then it is of little surprise that possessors of SQUARE can reliably recognize squareness when they see it.²¹⁴ Anyone who cannot recognize squareness, so the thinking goes, must either (1) lack a concept with the sort of structure that would lead them to draw the relevant inferences, or (2) possess a concept with structure which is deviant or defective in some way, i.e. a concept with a structure that disposes them to make the wrong inferences. For instance, suppose someone reliably says “square” when asked what shape a triangle is because the concept expressed by “square” for them has features such as THREE-SIDED as constituents.²¹⁵

In this case there are two possibilities. If it is a case like in (1), then the concept that they express via their use of “square” is not a SQUARE. If it is a case like in (2), then the concept they express via “square” is a SQUARE, though one that is defective insofar as it contains the feature THREE-SIDED as part of its cognitive content and represents squares inaccurately. Though defects of such an extreme degree are commonly thought as a sign, not of a defective square concept, but rather as a sign of a total lack of the concept.²¹⁶ Thus in a case of this sort,

²¹⁴ Peacocke (1992) is perhaps the most explicit of the philosophers in tying concept possession to the possessor’s inferential dispositions.

²¹⁵ It is unclear what exactly concepts in these deviant cases are representing, e.g. whether the concept deployed is a correctly deployed TRIANGLE incorrectly associated with “square” or if it’s a SQUARE with deviant or defective cognitive content. The answer will likely depend on both the correct theory of concept possession and that of reference determination for concepts. This issue will arise again in ch. 6.

²¹⁶ See Fodor (2004a) for discussion. See Prinz (2002), pg. 88, for an example where a large degree of variation between two peoples’ theories may preclude them being theories of the same thing, because at some point the differences become severe enough that communication and psychological generalization breakdown.

I suspect more philosophers (and perhaps, people in general) would be inclined towards the answer given by (1) than that given by (2). In general though, the stricter the normative demands one places on concept possession, and thus the tighter one makes the connection between concept possession and the possession of a concept with a particular structure, the more incoherent scenarios fitting (2) become.

These normative requirements play more subtle roles than explaining some of our attributional tendencies, however. First, as mentioned in chapter 1, requiring that concept possessors draw true inferences and/or correctly recognize the world-aspects their concepts represent may provide a basis for conceptual analysis, and thus a methodology for philosophy. If I cannot possess KNOWLEDGE without being able to recognize knowledge when I see it, then if I possess the concept then I should be able to correctly judge when thinkers have knowledge as opposed to mere belief, e.g. judging that thinkers in Gettier cases lack knowledge. I could then, via the consultation of such hypothetical scenarios, slowly refine my understanding of the structure of my concept via an investigation of the inferences it disposes me to make. Without normative demands on concept possession and the structure needed to explain such demands, such an activity would be unlikely to result in much more than an understanding of one's own, likely idiosyncratic, understanding of that world-aspect the concept represents.

A related issue is that, absent a robust theory of how the referent of a concept is determined, there seems to be no means of figuring out what referential kind a concept belongs to other than looking to whether or not it renders its possessor competent at particular tasks.²¹⁷

²¹⁷ Note that this is not the same as saying that such competencies are what *makes* a concept represent what it represents. One could deny that meeting these normative demands is necessary for having the relevant concepts while still holding that we have no grounds for attributing that concept except via our recognition of a thinker's competencies.

And such a theory of reference determination for concepts is something which, unfortunately, both philosophers and psychologists largely lack, despite a few promising proposals. For instance, if a subject reliably sorts triangular objects into one pile and square objects into another, we seemingly cannot say whether they are correctly applying TRIANGLE and SQUARE or systematically misapplying concepts like DIAMOND and RECTANGLE. Considerations such as this can often be seen in discussions of the concepts of animals and small children, where their inability to draw the “correct” inferences or correctly discriminate and categorize aspects of the world are taken as evidence that they lack the corresponding concepts.²¹⁸ If, however, possession of a concept is necessarily tied with competency in relevant tasks, and thus tied to possession of a concept with a particular structure that makes possible that competency, then the possibility of systematic misapplication of concepts is blocked or at least blunted, and with it worries of this kind concerning what exactly someone’s concepts actually represent.²¹⁹

It is a consequence of using structure to explain these normative demands that the capacity to think about some aspect of the world becomes tightly and inseparably connected with the capacity to both represent that aspect in a particular way and to engage with that world aspect productively.²²⁰ This, in turn, means that thinking about some specific aspect of the world requires one to have a concept with the right sort of structure, where “the right sort”

²¹⁸ See Neander (2006) for an example of this line of argument as applied to the classic debates surrounding toad visual perceptions. See Millikan (2000) for an argument against Neander’s position.

²¹⁹ This may, of course, be less a point in favor of tying concept possession to abilities and more a reason to be suspicious of any attempt to attribute concepts absent a robust semantic theory for concepts. This point, that developing a theory of concepts and an account of who/what possesses which concepts is impossible without first considering their semantics, is one that I will return to in ch. 6.

²²⁰ This is what I take to be the core tenet of what Fodor termed concept pragmatism: drawing necessary connections between the possession of certain practical capacities and the possession of corresponding concepts. See Fodor (2004a).

means structure which enables the concept's possessor to satisfy those demands, e.g. drawing true inference, correctly identifying things in the concepts extension, and so on. For instance, being able to think about water requires that one possess the corresponding concept, WATER. If the concept WATER is structured, then possession of the concept seemingly requires possession of a representation with the right structure. What sort of structure is "the right structure"? Typically, it has been taken to be that structure which enables the concept possessor to adequately engage with the relevant aspect of the world. Thus, the concept WATER would have structure which would enable its possessors to recognize water under normal conditions (however those are defined), infer that it has certain properties relevant to the possessors needs (that water can be imbibed, for instance), and so on.

In summary, there are three phenomena that conceptual structure has often been marshalled to explain:

1. **Core Cognitive Capacities:** Our capacities to categorize, reason about, and comprehend linguistic utterings and inscribings about aspects of the world.
2. **Cognitive Content:** Differences in cognitive significance between coextensive concepts, empty concepts, and differences in core cognitive capacities that are supported by coextensive concepts.
3. **Normative Demands:** Norms thought necessary for concept possession, the grounds for concept attribution, the possibility of conceptual analysis, and the manner by which we determine the referents of concepts.

Regarding (1), conceptual structure is posited to explain the processing underlying core cognitive capacities. The structure of concepts is what is exploited by such processes in the performance of relevant tasks, and absent such structure there is nothing for such processes to

work with and no basis for the possession of corresponding abilities. (2), conceptual structure is commonly taken to constitute the cognitive content of concepts. It is differences in the cognitive content of coextensive concepts, and thus differences in their structure, that accounts for differences in the cognitive significance of otherwise identical PA-supporting representations containing them as constituents. These differences, in turn, manifest as clear differences in behavior, in particular differences in how core cognitive abilities are exercised by possessors who fail to recognize they possess coextensive concepts of a single world-aspect. Finally (3), conceptual structure is often used to explain how there can be normative demands on concept possession. If concepts are structured, and if possessing concepts with such a structure confers certain abilities on its possessors, then possession of a concept will ensure that its possessor be able to engage with the represented aspect of the world in the relevant and correct ways. This, in turn, can help to explain how we attribute concepts to one another, how an analysis of concepts can lead to knowledge of the world a priori, and how we can determine the referent of a concept and guard against (to some degree) systematic misunderstanding of what a thinker is representing in thought.

Taken together then, these three phenomena seem to give ample reason to posit structured concepts. It is perhaps unsurprising then that conceptual molecularism is the predominant view in both philosophy and psychology and thus that most philosophers and psychologists have thought that concepts are structured entities of some kind. There appears to be much to gain and comparatively little to lose by endorsing molecularism. Motivating atomism then would seem to be a difficult task from the outset.

There is, however, an issue which molecularism faces that I will argue it struggles to solve without reducing the view to triviality. Both (1) expecting concepts to support normative

demands and in general tying the core cognitive capacities to what a concept represents, and (2) expecting concepts to support the capacity to represent the world in thought and thus support psychological explanations of both folk and scientific varieties, poses a problem in that both (1) and (2) are in tension with one another. The main reason for this is that the very properties that support the possibility of concepts fulfilling what's expected of them in (1) become an obstacle to their being the constituents of PA-supporting representations, and thus frustrate their fulfilling of (2). If the sort of structure that allows for (1) is held to be essential to the concept (which I will argue (1) requires, then concept publicity becomes endangered, as it is unlikely that all thinkers (especially those with false or inadequately robust beliefs about the world-aspect in question) have a concept with just the right sort of structure to meet the relevant demands. And concept publicity, I will argue, is a non-negotiable condition for the possibility of (2).

4.4 – What is Atomism?

As mentioned earlier, conceptual atomism is a broad approach to concepts originating with the work of Jerry Fodor which denies that, in some sense, concepts are structured.²²¹ Given the explanatory power that structured concepts seemingly allow for within both psychology and philosophy, the atomist position has appeared to many (*especially* to those in psychology) to be unpalatable or even entirely unviable. The theory, however, enjoys substantial support within philosophy. This support is also broad, with many of its philosophical supporters agreeing on little else.²²² Something attracts many philosophers to the atomist position, despite the clear appeal of molecularism. Part of my goal in this section is to

²²¹ Fodor (1998).

²²² Fodor (1998, 2008), Kwong (2007), Laurence & Margolis (1999), and Millikan (2017).

determine exactly what motivations there are for adopting atomism and which of these are or ought to be thought compelling.

4.4.1 – Proposal CA1: Concepts Lack Structure

Determining atomism’s merits and demerits, however, requires getting clear on what atomism is, and more importantly is not, claiming. While there is some variation in how the view is understood (especially with Fodor, the view’s originator and main proponent), the understanding of atomism as a straightforward denial of the existence of conceptual structure is common and will serve as a good starting point. Take as one prominent and representative example the statement of the view Fodor gives in his 1998 book on concepts:

“What is the structure of the concept DOG?” . . . on the evidence available, it’s reasonable to suppose that such mental representations have no structure; it’s reasonable to suppose that they are atoms.²²³

A small sampling of other philosophers’ statements on atomism conform to this simple characterization of the view. Atomism has been variously defined as the view that “lexical concepts are primitive . . . they have no structure,” that “concepts lack ‘internal’ or constituent structure,” that “lexical concepts are unstructured,” that “(almost) all lexical concepts are unstructured symbols.”²²⁴

This simple and common understanding of the view can be stated as follows:

CA1: Concepts have no structure.

²²³ Fodor (1998), pg. 22.

²²⁴ In order, Laurence & Margolis (1999), Kwong (2007), Daley (2010), and Prinz (2002), pg. 89.

While this understanding of conceptual atomism is indeed common, as stated CA1 is far too broad. Of particular concern is that the proposed statement of the view fails to specify what kind of structure concepts lack. It cannot be that concepts lack structure in all respects. For instance, if concepts are mental particulars realized by complex neurological states (as seems likely in the case of humans and all other animals we know of), then they are at least structured at the level of description the neuroscientist concerns themselves with. For instance, if a token concept is realized by a particular pattern of neuronal firing in the frontal cortex, that pattern is presumably articulable into parts, and with it the concept when appropriately described.²²⁵ Likewise it would seem that any thinking thing we may ever encounter is likely to have concepts structured in this very low-level way, whether that turns out to be a structure consisting of transistors and silicon, vibrating crystals, Chinese citizens with hand radios, or even the parts of an immaterial soul.

I take it that not many engaged in the debate over conceptual structure mean to deny that concepts are structured in ways such as these.²²⁶ Rather, the atomist must instead be denying that concepts have a particular sort of structure, i.e. that concepts are unstructured when suitably described. Thus, it is important to determine exactly what sort of structure the atomist is denying the existence of in order to understand what the view concerns.

²²⁵ This is, I suspect, an obvious point. It is, however, worth drawing attention to for the reason that many of the issues in the concepts literature turn on how many theoretically relevant types concepts fall under and the relations between these types. These types include possible structural types that concepts may fall under.

²²⁶ Those who may be theorists who deny (or at least do not endorse) that concepts are mental representations, either because they identify concepts with entities in another ontological category (Peacocke 1991) or believe that theories of concepts are not committed to any one position on what concepts are (Rosch 1978).

4.4.2 – Proposal CA2: Concepts Lack Structure Relevant to Psychology

One natural possibility is that the sort of structure the atomist has in mind is that structure describable in the terminology characteristic of psychology. Indeed, it is the psychologists and cognitive scientists more generally who have most often been the explicit target of the atomist, so the structure of concern to them would by extension be a plausible target for the atomist's ire.²²⁷ Psychologists do not frame their theories in terms of neural structures and processes. Rather, they frame them in terms of mental representations, mental states, their semantic and perhaps syntactic properties, and the relations between all the above and entities in other relevant domains.²²⁸

Thus, another possibility is that the atomist is claiming something like the following:

CA2: Concepts have no structure relevant to psychological theorizing.

This, however, is still too broad. Among those properties of a concept that are relevant to psychology, especially for theories that are broadly in the CTM tradition, are a concept's formal or syntactic properties. It is these properties, after all, that make a difference to processing under classic CTM explanations. Many atomists though are happy to talk of a concept's syntactic properties, and even to call such properties a concept's syntactic or formal structure.²²⁹ Indeed, they would have to hold that concepts have such syntactic structure if they are to think that concepts can be the sorts of things to factor in computational mental processes

²²⁷ Fodor (1998) makes this as explicit as possible by building it into the title of the book.

²²⁸ This is, of course, controversial. Whether psychology really is autonomous and can avoid reference to neurology, biology, ecology, and so on is a matter of deep and extensive debate.

²²⁹ Laurence & Margolis (1999).

in the first place.²³⁰ Further, some atomists rely upon the syntactic structure of concepts to explain phenomena that other theorists rely upon more traditional conceptual structure to account for.²³¹ Fodor, for example, denies the existence of cognitive content, and due to this relies instead on differences in syntactic structure or other formal properties between coextensive concepts to explain their differing cognitive significance.²³²

Even the atomist then does (and likely ought to) allow for the existence and explanatory importance of some conceptual structure that is of use to the psychologist. As such, atomism cannot plausibly be held to be the position that concepts lack all structure describable in psychological terms. Atomism must then be denying the existence of *some* kind of structure that is relevant to psychology, but not denying the existence of all such structure.

4.4.3 – Proposal CA3: Concepts Lack Representational Structure

A plausible suggestion is that the atomist is denying the existence of a kind of structure that not just psychologists, but philosophers as well, have long thought concepts must possess: representational or semantic structure. On this interpretation the atomist doesn't deny that concepts have structure altogether; rather, they deny that concepts have *meaningful* structure, structure which encodes information about the surrounding world and its many aspects. This seems to be the view that Prinz has in mind when he defines atomism as “the view that (almost)

²³⁰ This is assuming that they endorse a classical computationalist view according to which computational processes are sensitive only to the formal properties of a mental representation. This assumption, however, does not seem to be uncommon among atomists.

²³¹ Some go even further than this, holding that concepts lack semantics entirely and that all phenomena concepts are meant to explain can be explained by their syntactic structure alone. See Chomsky (2000) and McGilvray (1998).

²³² Fodor (1998). Here Fodor goes so far as to outright identify the formal properties of concepts with modes of presentation. See Aydede (1998) for an argument against such an approach.

all lexical concepts are unstructured symbols,” and defines a symbol as unstructured when “it has no components that are semantically interpretable.”²³³

Consider, for instance, a simple example from the prototype theory. On that view, a concept like GOLD can be decomposed into other contentful representations, such as YELLOW IN COLOR, METAL, HAVING THE ATOMIC NUMBER 79, and so on. Further, the explanation of the abilities such representations underly relies upon the semantic properties of these constituent representations. For instance, it is in part because GOLD represents gold and METAL represents metal, along with the fact that the latter forms part of the structure of the former, that possessors of the concept GOLD are disposed to infer that gold is a metal.²³⁴ The semantic structure of the prototype is necessary for giving the explanation of the inferential capacity in question. Perhaps then it is specifically representational structure of this sort and corresponding explanations that the atomist finds objectionable.

This understanding of conceptual atomism can be stated as follows:

CA3: Concepts have no representational structure.

While closer to the mark, even CA3 will not do. One reason is that the issues of most concern to the typical atomist do not demand anything as strong as CA3. The atomist can, I will argue, get all they want out of concepts without giving up on the explanatory gains representationally structured concepts make possible. Thus, even if the atomist is tempted towards something like CA3, they should not be.

²³³ Prinz (2002), pg. 89.

²³⁴ The explanation is, of course, much more elaborate than this. The exact nature of the prototype’s structure needs to be spelled out in detail along with the nature of the processes which exploit that structure. And as seen in ch. 2, there is no one single answer as to the nature of prototype structure and processes that has achieved consensus among psychologists.

Consider one prominent motivation for atomism, the problem of ignorance and error.²³⁵ If molecularism is right, then it seems that for someone to possess a particular concept they need not just a concept that represents the relevant aspect of the world but in addition a concept that represents it in the right (or at least some specified and determinate) manner.²³⁶ This poses a problem in cases where people lack much information at all about the world-aspects they intuitively can think about or where they have mostly false beliefs about these aspects. For the atomist, it would be a problem if anyone who failed to know that whales are mammals or who thought that whales are fish thereby failed to possess the concept WHALE and thus could not think about whales at all. Were that the case, then they could not have the false belief about whales in the first place. The atomist concerned with this problem should therefore demand that possession of a token concept with a deviant or impoverished structure not preclude that concept from belonging to the relevant referential type.²³⁷

Rectifying this problem does not require them, however, to demand that concepts lack representational structure entirely. Concepts can be possessed even by the ignorant or wildly confused provided that the connection between the structure of those concepts and their identity is severed. Worries about the stringency of concept possession on molecularist theories

²³⁵ See Laurence, Margolis (1999), pg. 21-23, for an overview of this issue, though the issue recurs throughout the article as they survey the major theories of concepts, nearly all of which struggle with the problem.

²³⁶ The possession of a concept with a particular referent and possession of a concept with the right structure may not be separable (see preceding discussion). Note that this is the case not just for internalist theories but for certain externalist theories as well. For instance, Peacocke's (1992) theory is externalist but nevertheless demands that concept possessors be disposed to make the inferences individuating of the concept.

²³⁷ Worries of this sort are hardly novel to the atomist. The problem of ignorance and error can be found in Burge (1979), Kripke (1980), and Putnam (1975) as well, though in those cases concerns the need for a word's referent to be independent of the word user's idiosyncratic understanding of said word.

are not motivated by their endorsement of conceptual structure per se. Rather, they are motivated by the further step molecularists are often tempted to make, namely tying a concept's identity and thus possession to a concept's being structured in a particular way.²³⁸

Consider the related issue of concept publicity. What publicity demands is that the type-identity of concept tokens across individuals remain the same despite variations in their beliefs, inferential dispositions, and the like. This, in turn, requires that it is possible for concept tokens across different people to be type-identical despite those concepts differing in their structure. After all, it is the differences in concept structure that accounts for differences in many of their beliefs, inferential dispositions, and so on.

Again though, nothing about this demand requires that concepts lack representational structure altogether. All it requires is that (again) it be possible for particular concepts to remain of the same concept type across structural variations. For instance, a concept like FRUIT is common across a wide variety of cultures and across individuals with drastically different levels of cognitive development and knowledge of fruit. Further, the ways in which individuals think about fruit will differ depending upon their practical needs: a chef and a botanist put fruits to different ends, and their knowledge and habitual ways of thinking of fruit differ accordingly.²³⁹ These differences are captured by differences in the structure of their FRUIT concepts, and these differences are in turn shown by real differences in their behavior. Botanists will readily apply FRUIT to tomatoes while chefs will not (and supreme court justices will apply the concept in a context-sensitive manner), and such a phenomenon is precisely the

²³⁸ As seen earlier in the chapter, this temptation is a motivated one, as making this tight connection does bring with it various benefits, even if the costs arguably outweigh them.

²³⁹ These issues are discussed at some length in ch. 3, in particular § 3.2.

sort that concepts were posited to explain. If FRUIT were individuated in terms of either the chef's or the botanist's cognitive content for FRUIT, i.e. the particular structure of either of their concepts, then one of the two of them does not really possess the concept, and thus lacks the capacity to represent fruit in thought at all.²⁴⁰

Similar considerations apply to concept stability, i.e. sameness of concepts within a single individual over time. Contemporary theories of cognitive development require that a concept's identity can be stable across the changes in structure that occur over the course of an individual's maturation. As before, this demand does not require that concepts be unstructured. It only requires, again, that whatever structure they have is not essential to them. For example, very young children tend to classify things according to superficial, observable properties. It is a major conceptual milestone in young children when they cease focusing on superficial characteristics and making purely similarity-based judgments and shift instead to categorizing on the basis of deep properties such as the animal's lineage or an artifact's intended function. This shift in the child's thinking and categorization behavior is explained in terms of a shift in the child's concepts, in particular as a change in the child's theory of biological kinds of which their animal concepts are constituents.²⁴¹ However, explaining such development in this manner requires that at least some of the child's concepts remain stable across time and that

²⁴⁰ In ch. 7 it will be discussed in detail why this is a serious problem for psychological theories, whether folk or scientific. In short, the issue is that it makes it impossible to explain misrepresentation, and in turn this makes providing explanations of how thinkers succeed or fail in their purposeful engagements with their environment impossible to formulate.

²⁴¹ The theory-theory of concepts is currently the dominant position in developmental psychology, so studies of conceptual development tend to be framed in terms of the development of the child's theories of the world. Note also that it is controversial exactly what concepts are on the theory-theory. Some hold that they are theory constituents, the mental equivalent of theoretical terms for scientific theories. Others, however, hold that they are theories themselves or theory-portions. I will be assuming they are theory constituents in what follows.

those concepts are possessed prior to full and complete development. The child is acquiring new ways of thinking about animals like raccoons as they develop, and this requires that they already possess concepts like ANIMAL and RACCOON early in development, acquiring not concepts of new things but new concepts of familiar things. It is because the child's already existent concept of animals shifts in structure past the age of five that their categorization behavior with regards to these animals differs from their younger peers.

A second issue with CA3 concerns the notion of structure itself. Remember how broadly structure was construed earlier. Structure includes not just the internal constituents of a concept (what was termed *internal* structure) but also those representations *external* to it and to which it stands in inferential, functional, or causal relations (external structure). The atomist can consistently and reasonably deny that concepts lack internal representational structure, though the motivation for such a claim is unclear given that it is unnecessary for what the atomist wants. It is, however, at least reasonably clear what such a denial amounts to. However, when it comes to external structure, atomists need not, should not, and indeed do not deny that concepts are so structured. Nor is it especially obvious what such a denial would amount to.

The conceptual atomist will, like every other theorist of concepts, need to say something about why typical concept possessors behave in the ways they do, e.g. that are inclined to draw certain inferences and not others. They will also need to say something about how concepts connect with other sorts of mental representations, such as perceptual or motor control representations. The explanations given for these phenomena, even in the case of the atomist, are likely to appeal to things like functional relations both between concepts and between concepts and other types of mental representation. Thus, provided that external structure is understood in a broad and unrestricted sense such that it encompasses any causal

and/or inferential relations of potential explanatory importance that the concept stands in, even the atomist will agree that concepts are externally structured and that such structure is representational in nature.

Granted, a molecularist will likely balk at the claim that the atomist can hold that concepts are externally structured. It is true, they may say, that if external structure is understood in an unrestricted sense, then an atomist's concepts will be externally structured. But "external structure" should not be understood in such a loose sense; this is not the sense of "external structure" that the molecularist has in mind. Rather, they think that it is only some subset of the explanatorily relevant causal, functional, or inferential relations the concept stands in properly count as its structure.

This complaint, however, shows that the debate between the molecularist and atomist is not really over whether or not concepts structured. There is no reason for the atomist to deny that concepts can have parts that are themselves representational, i.e. that they can have internal structure. Nor is there good reason for them to deny that concepts can stand in causal, functional, and inferential relations to one another and to other sorts of representations, i.e. that they can have external structure. Rather, it is a debate over the specific status of those structures or parts of them. It is a debate over whether those literal parts of a concept, if they exist, are essential to it. And it is a debate over whether we can draw an explanatorily relevant distinction between those representations standing in causal, functional, and inferential relations to the concept that properly counts as its structure and those that do not.

4.4.4 – Proposal CA4: Concepts Lack Individuative Representational Structure

What sort of special status may be had by those representations which are properly part of the concept's structure? What separates these representations from ones similarly related to

the concept, but which fail to be a part of the concept's structure? The natural possibility, and the one many participants in the debate have settled upon, is that such structure is individuating of concepts:

CA4: Concepts have no representational structure which is individuating of the concept.

The suggestion here is that it is not a concept's representational structure that makes it the kind of concept that it is. What makes the concept the kind of concept it is involves something different, such as perhaps what it refers to. The representational structure had by a concept is merely incidental to its identity if CA4 is true. This seems to be how Block understands the distinction between molecularist and atomist position:

Molecularism characterizes meaning and content in terms of relatively *small parts* of the web [of belief] in a way that allows many different theories to share those parts. For example, the meaning of 'chase' might be said by a molecularist to be *try to catch*. Atomism characterizes meaning and content in terms of *none* of the web; it says that sentences and beliefs have meaning or content independently of their relations to other sentences or beliefs.²⁴²

As with the other proposals before it, CA4 will not work. One reason is that on any atomist proposal a concept's structure, even if construed in a restricted manner, may still be allowed to play a role in individuating concepts. Further, given how broadly "structure" has been understood in much of the literature, many atomist proposals do, in fact, end up allowing structure to play such a role.

²⁴² Block (1998).

Take a toy version of Fodor's informational semantics, where a concept's referent (and thus the concept's identity) is determined in part by that concept being causally correlated with its referent (the niceties involved in specifying exactly which correlation is the relevant one need not concern us here).²⁴³ The concept's structure may be relevant to setting up and maintaining the causal correlation between concept and referent, in this way being relevant to determining the concept's identity. For example, one's concept WATER may represent water due to its activation being reliably caused by and thus correlated with water. But there is still the question of how WATER gets reliably activated in the presence of water. If it is, and this is not a matter of mere accident but rather a lawful occurrence (as Fodor requires), then there must be some causal mechanism or other attaching concept to world.

One possibility: the concept WATER may be reliably activated by the presence of water due to its inferential connections to concepts like CLEAR, ODORLESS, and LIQUID, or by its relation to perceptual representations of those properties. A concept may even get its content by being reliably linked to the world via the very representations psychologists theorize about, e.g. prototypes, exemplars, theories, etc. One's concept WATER may be activated reliably due to its connection with a prototype or a set of exemplars for water, or due to its place in a larger folk theory of chemical kinds. In such a case, the prototypes, exemplars, and theories would be acting as what Fodor terms sustaining mechanisms for WATER. Sustaining mechanisms are fittingly whatever mechanisms setup and sustain the reference-determining relation between the concept token and the aspect of the world that it represents. Thus, even on Fodor's view,

²⁴³ As stated, this view is woefully inadequate as a theory of semantic determination for concepts, running straight into problems like indeterminacy and accounting for coextensive and empty concepts. But a simple view such as this illustrates the idea just as well as using a more robust view such as Fodor's mature asymmetrical dependence theory, and does so without unnecessary digressions and exposition.

conceptual structure can affect concept individuation due to the influence such structure has on setting up and maintaining the referential connection between concept and world.

One complaint here may be that such sustaining mechanisms are not truly individuating of the concepts in question. Having a concept that is connected with a sustaining mechanism that relies upon representations like ODORLESS, CLEAR, and the like is not at all sufficient for that concept being a token of WATER. As Putnam long ago argued, the thinker in which this mechanism is housed must also be embedded in the right sort of environment, one in which the clear, odorless liquids around them are water and not XYZ, and related to it in the right manner.²⁴⁴ Were I on Twin-Earth, my sustaining mechanism that here sustains my concept WATER would instead sustain a concept XYX instead.

More importantly, it may be reasonably objected that this is not so much a case of conceptual structure individuating concepts à la molecularism, so much as it is conceptual structure determining or causing the properties of concepts (in this case, what they refer to) by which they are individuated. In other words, such structure looks to be more of an enabling condition for a concept meeting the conditions that individuate it, rather than being those latter conditions themselves. It may be true that a prototype or some loose set of representations inferentially linked to the concept is responsible for the concept bearing the right relation to what it represents. But it is standing in this relation, not being related to what sets up this relation, that makes the concept the kind of concept that it is. In the same manner, what makes

²⁴⁴ Putnam (1975). It is worth noting that this is a problem that also afflicts many of the less sophisticated versions of molecularism. It does not, however, affect versions of molecularism such as Peacocke's, where a concept's structure includes inferential relations to aspects of the environment, e.g. it's part of the essential structure of RED that its possessor is compelled to infer the presence of redness upon having a perceptual experience of redness.

an alloy bronze and is thus individuating of bronze is that the material consists primarily of copper with a small amount of tin mixed in. That this composition arises via blacksmiths and their use of forges does not, however, make blacksmithing individuating of bronze as an alloy.

A clearer case may be Millikan's recent unicept/unitracker theory.²⁴⁵ On this view, unicepts (internally unstructured concept-like representations) are individuated via their connections to their corresponding unitrackers, mental mechanisms (either perpetual in nature or portions of uniceptual networks) whose etiological function is to reidentify substances across differences in media and collect information regarding them. For instance, on this theory one's water unicept is a token of that type, as well as represents water, because its corresponding unitracker functions to reidentify and collect information about water, with this information encoded by the connections between the water unicept and other unicepts. There is no story to be given for how a unicept kind is individuated on this theory that does not appeal to the unicept's relation to other mental representations and mechanisms. In other words, it is not (as under Fodor's view) that the unitracker causes its corresponding unicept to have the property of representing what it represents. Rather, the unicept having that referent (and thus having the identity it has) is constituted by its being connected to a unitracker that tracks and collects information about the referent. The external structure of unicepts (either restricted or unrestricted) then plays for them an individuating role. Nevertheless, Millikan holds that her theory is an atomist one; unicepts are unstructured in whatever the relevant sense is.

It seems then that we cannot understand atomism (at least as its proponents do) in terms of whether conceptual structure is individuating of concepts, as on standard atomist proposals

²⁴⁵ Millikan (2017). See ch. 3 in particular.

structure can play a role in individuation. Being related to a sustaining mechanism or unitracker of the right sort is sufficient on many atomist views for a concept or concept-like representation to belong to a particular concept type. What is worse is that there are deeper reasons for why the atomist need not oppose the individuation of concepts in terms of their structure.

4.4.5 – Proposal CA5: Concepts Are Referentially Individuated at a High-Level of Abstraction and Multiply Realizable

What should the atomist focus on given their worries about such things as publicity, stability, and the ignorance and confusion of many concept possessors? The answer, I think, is implicit in the discussion throughout this chapter and has shown up several times before, most notably in chapter 3. Recall the examples of the chef and the botanist, the child’s developing ways of thinking about biological kinds, and examples involving competency with names such as “Einstein” and terms like “arthritis” familiar from the philosophy of language.

What I argued with such examples (and many others have argued elsewhere) was that concepts need to exhibit a certain sort of publicity across their possessors and stability across time (both stability within and without an individual thinker). Explanations of cognitive development depend on concepts maintaining their identity across the structural changes that development produces; otherwise any change to such structure would result in the replacement, rather than the augmentation or changing of, a pre-existing concept. Likewise explaining commonalities in behavior require sameness of concepts in some sense, despite differences in their concept’s associated structures. Were concepts not public in this way, psychological generalizations and explanations would break down.

Concepts thus need stable identities even in the face of potentially significant structural differences, both between distinct concepts and between multiple stages of a single concept.

As stated by Rey in his discussion of Machery's theory of concepts:

The problem is that "bodies of knowledge" vary between people and stages in a person's life. Unless one restricts the relevant knowledge in some principled way, no two people (or stages) will share a concept, since, short of coincidence, no two people (or stages) will bring exactly the same knowledge or procedures to bear in making many of their judgments.²⁴⁶

Talk of a concept's identity, or perhaps better its type, is however grossly ambiguous. And whether two "bodies of knowledge" count as the same concept depends on what sort of "sameness" one has in mind. As mentioned all the way back in chapter 1, everything belongs to a wide variety of types, and each of those types will have different principles of individuation. For instance, if someone were to ask, "How are automobiles individuated?" the answer is "It depends." We can individuate automobiles in terms of make (Ford, Chevy, Toyota), model (Focus, Camaro, Prius), body style (sedan, truck, minivan), condition (new, used), function (mere transportation, hauling cargo, providing emergency medical services), and so on. What distinguishes sedans from minivans is their general form, while what differentiates makes is the dictates of executives at major automakers. Each way of individuating automobiles gives us a different set of kinds and corresponding essential properties (it is essential to a used Prius that it had previously been bought and used by someone, but not to a Prius as such).

²⁴⁶ Rey (2010).

As Machery notes, this same point applies, not just to ordinary categories like cars and sedans, but to the kinds of the sciences as well:

Many entities can be legitimately individuated in several ways ... Organs such as hearts or eyes can be individuated in numerous ways, for instance, in a phylogenetic way and in a functional way. According to the first kind of individuation, two organisms have the same organ if and only if they possess the relevant organ by virtue of common descent. According to the second kind of individuation, two organisms have the same organ if and only if the relevant organ fulfills the same function. Depending on how organs are individuated, opposite answers might be given to the question “Do humans and cephalopods have eyes?”²⁴⁷

This exact point applies to concepts as well. “Are concepts individuated structurally?” Well, it depends on how you carve up the domain. There are at least three ways of individuating concepts that are relevant to the discussion here²⁴⁸:

Referentially: In terms of the concept’s referent, e.g. concept of cats, concept of dogs, etc.

Coarse Structure: In terms of the coarse structural properties corresponding to each of the major theories of concepts, e.g. prototype of a cat, exemplar of a cat, etc.

Fine Structure: In terms of the specific relations to other representations and/or constituent representations found in particular coarsely structured concepts, e.g.

²⁴⁷ Machery (2009), pg. 33.

²⁴⁸ There are, of course, ways of individuating them that are not relevant here but may still be important, such as in terms of their syntactic or neurological properties, among many others.

prototype that represents cats as being typically furry, prototype that represents purring as a cue-valid behavior for cats, etc.

Which of these families of types are relevant for the sort of publicity and stability the atomist demands? It is, I believe, only the referential types.²⁴⁹ For instance, for us to communicate with one another we must, at the very least, have the same things in mind as we converse. And this requires, in turn, that our thoughts contain type-identical concepts of the right, i.e. referential, sort. But if your thought consists of a prototype and mine an exemplar, there's little reason to think communication will break down due to this fact alone. A child with nothing more than a prototype of water can converse just fine with a chemist whose primary means of thinking of water is a theory: they will both refer to the same things as water in typical cases, they will both recognize paradigmatic instances of water as such, they will both perform the same on ordinary tasks that involve water, and so on. And there is certainly no reason to think that communication breaks down if my prototype of cats highlights their furriness and yours highlights their pointy ears.

Sameness of either coarse or fine structural type is not a prerequisite for successful communication, save for cases where these differences are severe enough that they block recognition of the correct referent in the relevant linguistic context. The child will, of course, fail to recognize some things as water due to their preoccupation with its superficial properties. And I may fail to realize you are talking about your sphinx cat when you use "cat" if my prototype overestimates the typicality of feline furriness. But these are mistakes people do in

²⁴⁹ Though as I will later argue, there may be forms of publicity other than this that are required for certain kinds of explanation. These sorts of explanations and concerns are not, however, of the sort that atomists are usually concerned with.

fact make and cases where communication may well break down. People do succeed in talking about the same world-aspects while sometimes simultaneously failing to recognize this fact, due to the odd and idiosyncratic ways they think about these aspects and the judgments these ways of thinking dispose them to make. Thus, they are phenomena which a theory of concepts should account for. These same considerations apply *mutatis mutandis* to issues regarding the stability of concepts across conceptual development and concept possession by the ignorant and confused.

Psychological explanations likewise require that we share concepts, but do not in the normal case require that those concepts we share are identical in either coarse or fine structure. It may be a true generalization about humans that as social animals we normally share a strong desire for romantic love and companionship, and thus must share a concept LOVE that factors in the representations which support those desires. But sharing that desire is not blocked if I have as part of my concept of love that sexual desire is inessential to romantic love and your concept has as a part that it is essential. We both desire the same thing; we merely disagree on what exactly that thing we desire is like. And it may be that humans pursue such relationships in part because they believe that it is an essential part of a good life for humans, though that hardly requires that their concepts of a good life agree in how they represent such a life to be.

Similarly, explanations of a child's cognitive development demand only that a child's concepts represent the same things across that development. A child's theory of biological kinds must change in its fine structure over time to account for the ways in which the child's reasoning changes across development. It is, after all, those changes in fine structure that are meant to explain how that development proceeds and why the child's behavior changes across

development. But that does not mean that their theory needs to change with regards to the world-aspects it represents in addition to the way it represents them as being.

What I propose is that the type-identity of concepts qua referentially individuated does not imply anything about the sameness of concept types at levels more finely grained, in the same way that you and I both driving the same car qua sedan implies nothing about whether we both drive brand-new navy-blue Ford Focuses. Such an implication *could* have been there; internalism theories and molecularist theories like Peacocke's that tightly bind together what a concept represents and how they represent it, if true, could perhaps support it. The atomist need not deny that this connection *could* have been there, at least contingently, in the same manner that the identity between sedans and brand-new navy-blue Ford Focuses follows in a world where Ford has a permanent and total monopoly on sedans and offers nothing in the way of product variety.²⁵⁰ Provided that there is, as a matter of empirical fact:

1. A level of concept types characterized by (a) being a representation with the function of factoring in processes supporting core cognitive abilities, and (b) representing some aspect of the world (however it comes to do this and in whatever way it does), and
2. These concept types are multiply realizable by a wide variety of concepts types individuated by their coarse structure (and these realized by many concept types individuated by their fine structure),²⁵¹

²⁵⁰ Perhaps there are worlds with nothing but creatures possessing incredibly limited representational abilities, for whom there is one and only one route to representing all there is for them to represent. Likely our own world was such a world during the earliest moments of the emergence of intelligent life.

²⁵¹ This is similar to Millikan's (2000) view that concepts are abilities insofar as she distinguishes between (1) individuating abilities in terms of their ends, e.g. the ability to navigate to some place, versus (2) individuating abilities in terms of their means, e.g. the ability to navigate via maps versus landmarks.

the atomist gets everything they need.

This understanding of atomism can be stated simply as follows:

CA5: Concepts are (1) individuated at the level of reference and (2) can (in the normal case) be realized by a wide variety of structured representations.²⁵²

This proposal can be visualized as in figure 4.1 below, with the parentheses in each box indicating what is relevant to the individuation of concepts on that level.

It is worth noting that this is somewhat of an oversimplification. For one, the distinction between coarse and fine structures is not absolute but rather a matter of degree. There are, for instance, a wide variety of different types of prototype representation, differing according to how complex their structures are and what kinds of information they encode, e.g. typical versus cue-valid properties, or prototypes that draw a distinction between a thing's attributes and features. There are thus likely to be ways of dividing up concepts that are coarser than in terms of what I have termed "fine structures," yet finer than in terms of what I have termed "coarse structures." I have simply selected what I take to be the most salient levels of concept types for the points I am trying to make.

²⁵² See Edwards (2011) and Lalumera (2010) for other reasons to think concepts are multiply realizable in the sense given here.

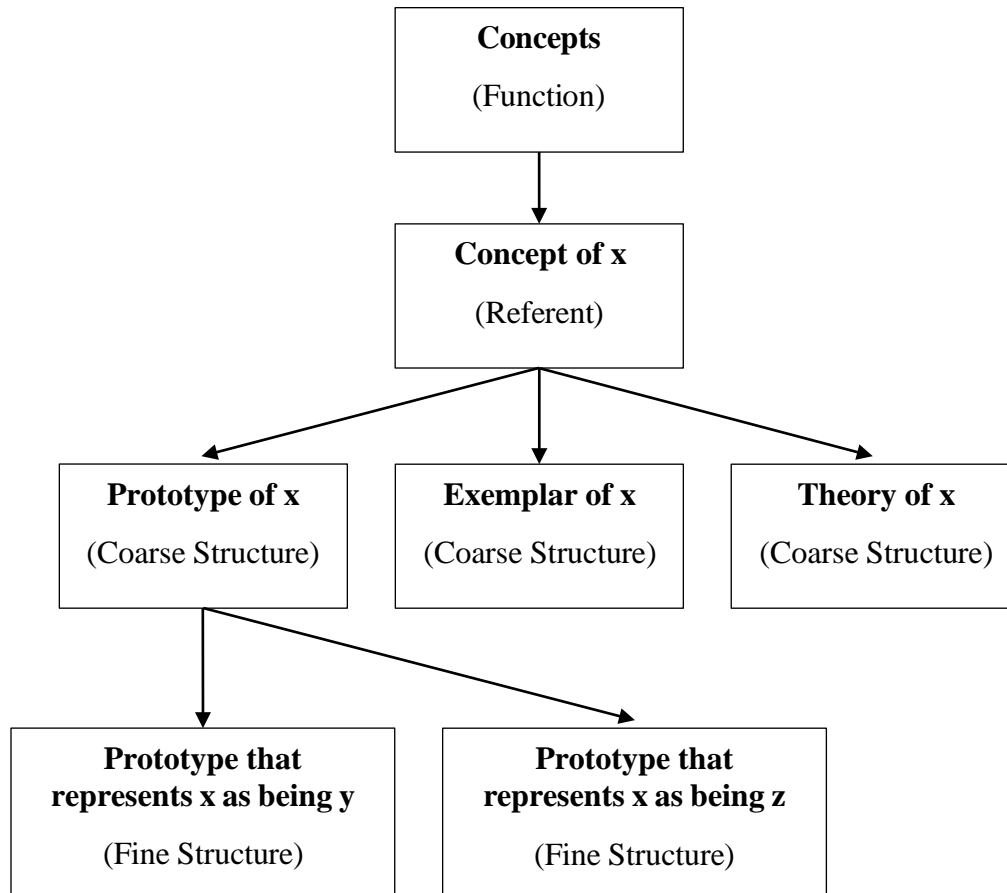


Figure 4.1

How exactly does this proposal differ from the molecularist proposals of philosophers such as Peacocke? The difference is twofold. First, for any (or nearly any), concept there are a wide variety of coarse structures that could realize it. A cat concept may be a cat prototype, a cat exemplar, a cat ideal, and so on. Second, for each of these coarse structural types, there are a wide variety of finer structures that could realize them. Many of us reliably recognize water on the basis of a prototype representing its typical visually perceptible properties. However, there is no necessity in this. Hellen Keller could not, for instance, token WATER on this basis.²⁵³

²⁵³ This example is a favorite of Millikan, though she does not use it in defense of atomism as she understands it. It does, however, serve to illustrate what I am here taking to be the central position of atomism.

Rather, she would have done so on the basis of representations of its tactile properties, or through representations of others communicating its presence to her via selectively applied pressure on her hand. A chemist could, in principle, know nothing of the observable properties of water, but theorize about, detect, and manipulate it based upon their theoretical understanding of its chemical properties alone. Finally, one could think of water on the basis of nothing more than the fact that one is embedded in a linguistic community where “water” is in common use, despite total (or near to it) ignorance of the properties of water. In other words, atomism of the sort defended here is compatible with the sort of conceptual pluralism I advocated for in chapter 3 in a way that standard molecularist proposals are not.

There is nothing strange about the sort of cases discussed above and in earlier chapters; they are not cases that exist only in the seminar rooms of philosophy departments. Thinkers rely on bespoke properties of things all the time to categorize them and often draw various inferences that few (if any) others would have reason to draw. Adapting an example from Millikan, I may regularly infer the presence of my housemate by nothing more than that their car is in the drive when I get home or the presence of my office mate by noting that the door to the office is unlocked when I go to open it. I regularly categorize bottles of water as such because they come in clear and distinctively shaped cylindrical bottles, and not in the soft pink/purple tapered bottles of nail polish remover or the boxy, brown bottles of hydrogen peroxide. And coming to think of something merely via exposure to a word for it is equally an everyday occurrence. This is, after all, the position that all children find themselves in when they wonder about the adult matters that they overhear their parents speaking of, or of the confused student who missed a previous lecture and now wonders what half the things being spoken of in class are.

The atomist then differs from the molecularist as understood here in denying that there is any one structure, any one way of representing something, which is essential to the possession of a concept and thus the capacity to think about that aspect of the world. While the molecularist holds that a concept's structure both partially individuates it and is essential to it, the atomist denies the latter claim and accepts the former only with major qualifications. To embrace conceptual atomism of this sort is to embrace a highly pluralistic account of conceptual structure and with it a highly pluralistic account of concept possession. Individuals with widely different ways of thinking and reasoning about aspects of the world can nevertheless all possess the same concepts and think the same thoughts, at least in one important and central sense of "same concept" and "same thought."

It may even be that the difference between molecularist and atomist is more a matter of degree than anything else, a matter of just how tight they take the connection between possessing a concept of some world-aspect and possessing a concept that represents that aspect as being some particular way. For instance, those philosophers who allow for some limited degree of conceptual variation may fall somewhere in the middle between full-throated atomism of the Fodor variety and the highly demanding molecularism of Peacocke. Perhaps there is only a very limited range of structurally individuated concept kinds corresponding to each referential kind, say a folk concept and an expert concept or that concept had by young children and that had by adults. While possible, such a view I think fails to appreciate just how varied real-world concepts are. Further, as I will argue in chapters 6 and 7, I think it reflects a mistaken view of how concepts come to represent whatever aspect of the world they represent.

4.5 – Some Potential problems

I will end on two worries that are, I think, natural to have about my proposal here. First, there is the worry that I have twisted atomism beyond recognition. We have moved step-by-step from atomism understood as the view that concepts are unstructured entities of some kind to atomism understood as the view that there is a level of concept types individuated referentially and a multiple realizability claim about such types. The second worry is, in short, that understanding atomism in this way is trivializing the position, making it into something close to a truism and thus depriving it of all that made it philosophically interesting.

Regarding the first worry, it may well be that what I call “atomism” is rather distant from what others have called by that name. However, it maintains what is important, namely a tight focus on accommodating the phenomena which have been of most concern to atomists. It safeguards concept publicity and stability, renders a non-issue the problem of ignorance and error, and I suspect (though will not argue) that it does not run afoul of Fodor’s concerns about compositionality and thus productivity and systematicity. It may be that it is consistent with views that have been described as molecularist by some theorists. But this latter issue, the potential compatibility between atomism and molecularism suitably understood, is a feature rather than a bug. If clarifying what exactly it is that atomists care about serves to bring self-described atomists and molecularists into agreement, then all the better.

Regarding the worry about triviality, what I have said here will, I think, be a truism to anyone who has fully internalized all the lessons of semantic externalism from the 70s onward and who takes the general truth of theories of concepts from scientific psychology as a datum. This all will not, I suspect, be revelatory for anyone who recognizes the need for minimal concept possession conditions for the much the same reasons Kripke, Putnam, and Burge

(among many others) stressed minimal conditions for competency in the use of words in a natural language. Nor does anything I say here in any way pose a threat to the psychologist and their theories (save for their internalist component, an inessential piece of those theories).

My target, those who do run afoul of atomism as I understand it, are all those like Peacocke who tie concept possession and conceptual competency tightly together. And if atomism as I understand it is true, then granted what seem to be empirical truths about concepts (such as that the young child thinks that painted raccoons are skunks, or that some adults think whales are fish), hopes such as Peacocke's and that of other philosophers, that concepts and their analysis can provide the philosopher with knowledge of the world from the comfort of one's armchair, are at best misguided. A child the age of five would be unwise to think they can learn about skunks from their consideration of hypothetical scenarios involving painted raccoons and what their concepts dispose them to judge. The philosopher should not, I think, take themselves to be much more secure on this front. And without that possibility goes one of the major motivations for molecularism as understood here.

Talk of "the" concept of something remains common in philosophy and has persisted long past the point many thought better of conceptual analysis as a philosophical methodology. If atomism as I understand it is correct, then the most that can be said about "the" concept of some world-aspect is that (1) it essentially represents that aspect, and (2) that it essentially plays whatever functional role is characteristic of concepts. Everything else that may be said about any given concept of that aspect, such as what properties of it are represented, how the representations of those properties are related within the concept, and even whether that aspect has any of the properties the concept represents it as having, is a wide-open question. Talk of some truth being conceptual or something's being conceptually necessary becomes, if not

nothing, then nothing of much philosophical interest. And if analyticity is a matter of conceptual containment, as Kant and some since him have thought, then analytic “truths” need not even be true, let alone necessarily so.

5 – From Concepts to Conceptual Atoms and Structures

In the previous chapter I argued that atomism should not be understood as the seemingly simple view that concepts are unstructured. This was due to three issues with such an interpretation: (1) the difficulties in spelling out exactly what it means for concepts to be unstructured, and (2) the limited explanatory power wholly unstructured concepts would offer, coupled with the desire of many atomists to accommodate psychological explananda, and (3) that such an understanding commits atomists to a stronger claim than they need in order to secure the things they want, such as publicity and stability for concepts.

Instead, I argued that the atomism of many philosophers today ought to be understood as the view that concepts are individuated (on at least some level) referentially, and that such high-level concept types are not necessarily realized by any one structured representation (or limited range of such representations) particular to them.²⁵⁴ On this view, atomist concepts are concepts by virtue of their function, i.e. that they factor in the processes underlying core cognitive capacities, and at a high level of abstraction are the type of concept that they are by virtue of what it is that they represent. On lower, more concrete levels, concepts are individuated as well by their coarse structures, resulting in coarse structural concept types like the prototypes, exemplars, theories, and the like posited by psychologists. And atomism so understood holds that concepts are further individuated at even lower, more fine-grained levels in terms of their fine structures, resulting in prototypes, exemplars, and the like with determinate cognitive contents. Finally, it leaves open as well the possibility that concepts are

²⁵⁴ While presented as two distinct claims about concepts, these two may not be separable. Holding that concepts are individuated in certain ways seems to entail their multiple realizability, as does their multiple realizability entail certain ways of being individuated.

individuated in terms of yet other properties not mentioned here when explanations appeal to these kinds.²⁵⁵

This atomist proposal, what I will call *high-level atomism*, can be summarized as follows²⁵⁶:

High-Level Atomism: Concepts are typed referentially but have multiple subtypes individuated by their coarse structures, and these structural subtypes in turn have multiple subtypes individuated in terms of their fine structures.²⁵⁷

Atomism on this way of thinking is an inherently pluralistic position, one stressing that a concept of some world-aspect can take many forms, and thus that there are many possible routes to and ways of thinking about it. It decouples the capacity to represent something in thought from actually representing it in any one specific way or limited range of ways. Atomism so understood then is opposed to the view that concepts essentially involve any one way or limited range of ways of representing their referents, e.g. the view that a concept is a

²⁵⁵ Some possibilities include properties relevant to psychological theorizing, such as the syntactic and neurological properties of concepts in general, or the property of being widely possessed during specific stages of cognitive development had by some finely structured concepts in particular. Other properties could include those that may be of interest to fields like anthropology or the social sciences more generally, such as the property of being widely possessed or absent entirely within a particular community.

²⁵⁶ This proposal is perhaps best understood, not as a simple exegesis of existing atomist works, but rather a statement of what atomists, given their explanatory goals and other theoretical commitments with regards to concepts, ought to adopt. While I will argue here that it does not go far enough, it does serve as a useful rhetorical waystation along the path to a more revisionary, and I hope fruitful, position.

²⁵⁷ Another way of stating high-level atomism would be as the view that at a sufficiently high level of abstraction concepts have their referents essentially and have their structures, whether fine or coarse, inessentially. Concepts typed more concretely do, however, have their structures essentially. I take it that if a type of thing is individuated in terms of a given property, tokens of that type have that property essentially, failing to be tokens of that type if that property is not possessed.

token of CAT not just because it (1) represents cats, but in addition (2) has a particular structure in virtue of which it represents cats as having specific properties, e.g. being typically four-legged, furry mammals.²⁵⁸

Atomism is not then so simple as saying that concepts are unstructured full-stop. Rather, concepts are unstructured only from a suitably abstract perspective, one focusing solely on their functional and referential properties, among possible others. It is only when one looks at the concrete implementations of such higher-level concepts, i.e. at concepts belonging to the lower-level structural subtypes that realize them, that the structure of concepts comes into view. An atomism of this sort, while unlikely to be satisfying to some atomists, is congenial to atomism as understood by its more moderate proponents.²⁵⁹

What I will argue now is that high-level atomism does not go quite far enough, and that some of the tensions it was designed to resolve persist. Instead of endorsing an atomistic view of this sort, I will argue that we ought to take the core lessons of the view and incorporate them into a new framework for thinking about concepts. My point in what follows will not be that atomism as understood here and in the previous chapter is incoherent or unworkable. My point will simply be that there are ways of carving up the theoretical terrain that may better aid the understanding. And the way of carving I propose will, I suspect, be conducive to more

²⁵⁸ On internalist views, (1) and (2) are not easily separable in this way; on such views it is in virtue of having a particular structure that the concept represents what it represents, and provided the relevant sameness of their environment, differences in what is represented between two thinkers must involve corresponding structural differences in the concepts they possess.

²⁵⁹ See Kwong (2007), Laurence & Margolis (1999), and Millikan (2017) for three representative positions.

frictionless theorizing and more productive discussions about concepts, both within philosophy and between philosophy and other fields like psychology.²⁶⁰

5.1 – Conceptual Atoms and Conceptual Structures

A view such as high-level atomism works mostly without incident if you restrict your focus to concepts as understood by psychologists and most in cognitive science.²⁶¹ The lack of structure is the main stumbling block to atomist concepts explaining the sorts of capacities psychologists expect them to explain. And there is little reason to think that such structure must be essential to do this sort of explanatory work.

It is a far less natural fit, however, if you want a theory of concepts that accommodates not just the explananda and desiderata of psychologists but those of philosophers as well. As discussed all the way back in chapter 1, concepts are expected not just to help explain how people categorize, reason, and engage effectively in other such practical activities. They are also expected to act as the constituents of PA-supporting representations, i.e. the constituents of the representations underlying and making possible our having of propositional attitudes. In so doing they are expected to also explain our capacity to represent the world and its many aspects in thought. They are sometimes expected to do a medley of other things as well, from explaining normative demands on concept possessors to providing the ground for a philosophical methodology founded on conceptual analysis.²⁶² It is far from obvious that

²⁶⁰ In could, in this way, be seen as a largely pragmatic proposal, one along the lines of Machery's (2009) advocacy for the elimination "concept" from our theoretical discourse to facilitate better and more productive theorizing.

²⁶¹ This is somewhat ironic in that its original proponent, Fodor, appeared to have little to no interest in accommodating the explananda of many psychologists studying concepts.

²⁶² Though as argued earlier, these are likely to be unreasonable demands to make.

traditional atomist concepts can meet these additional demands on them, even if they are allowed to be structured at suitably concrete levels.²⁶³

It is, I suspect, too great an ask for concepts to do all that has been expected of them by both philosophers and psychologists. Much has already been said regarding how to reconcile the tensions between these two understandings of concepts, without, I think, much in the way of satisfying answers. Certainly there has been a lack of any answers that have achieved widespread support between philosophers and psychologists both. Psychologists remain suspicious of atomism and some philosophers even remain hesitant to call prototypes and the like concepts at all.²⁶⁴ And rendering conceptual atoms fit to play the roles psychological theories demand necessitates distortions to the original idea underlying atomism, distortions which while defensible are far from ideal. Worse is that views of concepts that expect too much of them fall prey to other problems I will discuss in what follows.

The position I will propose here aims to resolve these issues, or more conservatively, to at least ameliorate them somewhat. It is, I suspect, not wholly original, though I have not yet seen it stated anywhere explicitly. The position's core idea is that there is no one kind of representation that acts both as a store of information to be exploited by our cognitive processes and also simultaneously as a building block for thought. In other words, there is no one kind of representation that fits both the philosopher's and the psychologist's understanding of "concept." If concepts are understood to be some representation that will conform to the

²⁶³ This is because, in part, such structure can only meet such demands if it is essential and there is limited, if any, conceptual pluralism. These issues are discussed in some detail in ch. 3 and 4.

²⁶⁴ Rey (1983), Peacocke (1992), Fodor (1998), and Millikan (2000), pg. 49.

expectations of both philosophers and psychologists, then to put it plainly there simply is no such thing as a concept. No one kind of representation plays both roles.

Both of these roles must, however, be played by something. Our more complex conceptual representations must be built up out of simpler elements in order to account for seeming properties of thought such as its productivity, i.e. the capacity to potentially form an infinite number of thoughts on the basis of a finite number of elements. And absent conceptual representations that act as stores of information in memory, any thinking thing would be rudderless in the execution of their cognitive tasks. That both roles require filling does not, however, require that a single sort of thing fill both.

Importantly though, both roles still seem to be related to one another. Without the capacity to represent something in thought, there could be no possibility of forming representations of that something which could serve practical ends. Thus, the representations that serve the role the philosopher concerns themselves with appear to be prior to and necessary for the possession of the representations that play the psychologist's role. And one possible explanation for the intimate relationship between the two is simply that the representations that philosophers call "concepts" are the constituents of the representations that psychologists call "concepts." Without the conceptual building blocks talked of by philosophers there would be nothing out of which to assemble the complex conceptual representations talked of by psychologists.

Representations of the first sort, the simple building blocks for all thought, I will call *conceptual atoms*. Representations of the second sort, the complex representations underlying

core cognitive capacities, I will call *conceptual structures*.²⁶⁵ For now, all that needs to be borne in mind is that conceptual atoms play the roles philosophers have expected of concepts, conceptual structures those roles expected by psychologists, and the former are the constituents of the latter. Later on in § 5.3 I will present my theory of how these two types of representation relate to one another and their respective roles in more depth.

5.2 – Two Arguments Against Traditional Concepts

In this section I will present two arguments that are indirectly against high-level atomism. More directly, these arguments are against a general picture of concepts that high-level atomism and many other theories of concepts assume. This general picture is one according to which:

1. Concepts play the functional roles expected by both psychologists and philosophers, being both (a) constituents of those representations supporting beliefs and other PAs, and (b) information storing representations to be exploited in the exercise of one's cognitive capacities.
2. Concepts are both coarsely and finely structured, such structure is required for their performing (b), and at suitably concrete levels concept types are individuated by structures of both sorts.
3. Following from (1) and (2), the constituents of our PA-supporting representations are themselves mostly complex, structured representations.²⁶⁶

²⁶⁵ Talk of conceptual atoms is taken from Fodor's atomism and related proposals. The term "conceptual structure" is taken from Jackendoff's work in linguistics (1990, 2011).

²⁶⁶ The "mostly" is because such a view of concepts might endorse a level of primitive concepts or perceptual representations out of which all concepts are ultimately built.

The result of endorsing this traditional picture, in particular the result of endorsing (3), is that conceptual systems end up bifurcated into two distinct representational levels. One level of representation is that which is represented by the concept itself. And the other is that level of representation that encompasses whatever is represented by the more complex representations these concepts combine to form. These representational levels (or perhaps more accurately, the sets of PAs that they support) I will, following tradition, term *conceptual* and *background knowledge* respectively.²⁶⁷

Before discussing the two arguments, it is worth noting first that the traditional view of concepts given here already runs into a problem in endorsing this distinction. As discussed back in chapter 3, there are empirical reasons to deny that such a distinction exists. As such, there are reasons to deny that conceptual systems are, at least for minds like ours, bifurcated into two levels in this manner. I will not, however, go into this issue in more detail here and will instead recommend consulting § 3.4 for more information.

The first of these arguments concerns structural individuation of conceptual representations and the explanatory value of the resulting concept kinds. The traditional picture of concepts places a great deal of importance on the fine structures of concepts, and traditional atomism still endorses the individuation of concepts (on an appropriately concrete level) in terms of them. However, I will argue that if concepts are regularly individuated in this manner, then given a robust concept pluralism there will end up being a vast proliferation of

²⁶⁷ The usual caveats apply. “Knowledge” is used by psychologists and even some philosophers in the concepts literature to mean something more like belief, with no requirement that what one “knows” even be true. In addition, the terms “background knowledge” and “conceptual knowledge” I take here to be equivalent to distinctions such as the dictionary and encyclopedic knowledge distinction in linguistics.

explanatorily impotent concept types. What this shows, I'll argue, is not that concepts are never productively individuated by their fine structures or that types characterized in fine structural terms will never be of explanatory value. Rather, the claim is one concerning emphasis, that the traditional view gives to such types too central a place and promotes a view according to which such concept types are the norm, rather than uncommon occurrences. It needs to be recognized, in short, that only under specific conditions will concept kinds characterized by a common fine structure constitute a natural kind, one which can serve a robust explanatory role in scientific psychology.²⁶⁸ Failure to do so risks elevating the fine structures of concepts to a privileged place and to projecting their structure on concepts of the same referential kind.

The second argument is that the functional role of concepts as understood by psychologists looks far closer to that had by beliefs (or the representations supporting them) than that had by the concepts of philosophers. Representations such as prototypes do not appear to behave as thought constituents at all, e.g. they do not seem to combine together with other prototypes to form, among other things, belief-supporting representations. Rather, they appear to function in ways similar to beliefs or belief-supporting representations themselves.²⁶⁹ Representations such as prototypes store information about one's environment to be later exploited by the various processes underlying thought and action. It may even be, as has long been argued, that not only do prototypes not compose to form such representations, but that

²⁶⁸ This emphasis is, I suspect, a holdover from the kind of molecularism considered in the previous chapter. Were concepts individuated by their fine structures, then there would be good reason to be interested in such structure, for the variety of reasons discussed up until now, e.g. explaining normative demands, allowing for conceptual analysis.

²⁶⁹ Or more accurately, perhaps, systematically related collections of such beliefs or belief-supporting representations.

they cannot, and thus cannot play the role of PA-supporting representations at all.²⁷⁰ This disconnect in function between the concepts of philosophers and those of psychologists gives, I will argue, reason to suspect they are not theorizing about the same thing.

Neither of these arguments are intended to nor are thought to be knockdown arguments; each simply brings to light some of the complications and problems that the traditional picture of concepts generates. I take it that it would be a point in favor of my own view if it manages to avoid generating these issues and others like them.

5.2.1 – The Unneeded Proliferation of Concept Kinds

Back in chapter 3, I argued that we ought to adopt a highly pluralistic theory of concepts, one recognizing that a wide variety of concept types exist. Among these are concept types defined in terms of a wide variety of coarse and fine conceptual structures. And in chapter 4, I argued that we need to recognize that abstract concept types individuated referentially can and are realized by a wide variety of concept types defined by their coarse structures, and these in turn are realized by a wide variety of concepts defined by their fine structures. Concepts thus divide into coarsely structured concepts including prototypes, exemplars, theories, definitions, and ideals (among others). And these further divide into an even wider array of finely structured concepts, such as prototypes as a kind dividing into a variety of prototype subkinds each representing whatever it is they represent as having different typical/cue-valid properties.

It follows from this that two concepts being type-identical at a higher, more abstract level does not, on its own, imply anything about their identity at lower, more concrete levels.

²⁷⁰ Fodor (1998), ch. 5. For my purposes here, only the weaker claim that such representations do not in fact compose is needed. However, that my view manages to sidestep the longstanding debates concerning the conceptual structures posited by psychologists and issues of compositionality is, I believe, a point in its favor.

Adopting a pluralistic theory of concepts and the corresponding conceptual taxonomy I recommend is necessary to accommodate what seems to be the empirical fact that concepts of the same referential type are highly variable between and within their possessors, both in terms of their coarse and their fine structures.

So far, all of this can be accommodated by the general picture of concepts underlying positions like high-level atomism. This is unsurprising, given that atomism so understood was designed to accommodate, among other things, the wide variability of concepts. There is, however, still a tension in the view, one concerning the structural individuation of concepts given such widespread conceptual variability. Specifically, such variability, I will argue, undercuts a major motivation to individuate concepts structurally in the first place.

Consider first variation in coarse structure. Concepts being individuated in terms of coarse structures results in the familiar prototypes, exemplars, and theories of psychology, among other representations. The need to individuate in terms of coarse structures is, I take it, clear; prototypes and their brethren are well-entrenched in psychology and have proven to be explanatorily fecund kinds. The theories corresponding to each type are by now decades old and each is well-supported by evidence from a variety of fields.

What is not as immediately clear is whether, and if so why, concepts should ever be individuated by their fine structures. That I *could* individuate this domain in this manner hardly shows that I *should* do so. I could individuate buildings by the eye color of their architects, but the resulting types would have presumably dubious importance to the field of architecture. If concepts are to be individuated in terms of their fine structures, there should be some explanatory need to individuate in that manner.

If conceptual variability in fine structure is widespread, then individuating concepts in terms of their fine structures will result in a vast proliferation of concept types. There will be as many types of the concept of some world-aspect *W* as there are different ways of representing *W*, or perhaps even different ways of weighting the ways that *W* is represented as being.²⁷¹ Further, these concept types will in many cases have very few instances (in some cases maybe even only a single instance in a single person). And some of these concepts will not be stable or long-lasting, such as ad hoc concepts built for a single occasion and then discarded, e.g. my concept of my drink left on a counter at a crowded party, or that of my hand during a particular game of poker (provided that hand is not so remarkable as to be committed to memory).

5.2.2 – The Epistemic Theory of Natural Kinds

One way of better making sense of this issue is to consider it from the perspective of natural versus non-natural kinds. On any theory of natural kinds (or at least all that I know of), these highly idiosyncratic concept types will fail to constitute natural kinds. And there are many reasons philosophers will deny such kinds that status, including their lack of explanatory value, their arbitrariness, their relatively few instances, and so on. I will, however, be relying on and largely assuming the epistemic theory of natural kinds throughout the following discussion.

²⁷¹ I assume here and throughout that differences in the weighting of features in representations such as prototypes and exemplars count as differences in the fine structure of such concepts. If owing to environmental differences your prototype of a private dwelling represents its being a house as highly typical, and mine represents such dwellings being houses as highly atypical (perhaps instead representing their being an apartment as highly typical), we not only seem to represent private dwellings as a kind differently, but would seem to have differing beliefs on the what is typical for dwellings as well.

What distinguishes natural kinds from non-natural kinds on this view is their degree of inductive potential, where this is understood in terms of the number of systematic inductive generalizations that the kind supports. The greater the number of possible inductive generalizations, the greater the inductive potential. This inductive potential, in turn, depends on the kind exhibiting a kind of stable, grounded uniformity.

Throughout I will understand the epistemic theory of natural kinds as committed to four claims regarding such kinds:

The Epistemic Theory of Natural Kinds:

1. Natural kinds are *uniform*, i.e. have members that share a substantial number of properties with one another.
2. Natural kinds are *stable*, i.e. the similarity in properties between members persists over time.
3. Natural kinds are *stable and uniform in a grounded way*, i.e. there is some underlying reason or mechanism accounting for the stability and uniformity of the kind.
4. Because of (1), (2), and (3), natural kinds *afford a high degree of inductive potential*, i.e. one can draw a substantial number of systematic inductive generalizations regarding such kinds.

For example, dogs form a natural kind on this view because there is a good deal you can learn about dogs via induction. You can learn that they look, sound, feel, or even taste a certain way. You can learn that they behave in various characteristic ways. And so on. This learning is possible because dogs are all similar to one another in a great many respects (the

kind is uniform) and the kind does not change drastically across time (the kind is stable). Were it the case dogs shared little in common with one another, there would be little to learn about dogs *as such*. And if dogs changed most of their properties from day to day, or the species regularly underwent rapid mutation, there would not be much you could know about them for very long. This similarity between dogs is due to them being reproductively bound together as a kind, i.e. they are each partial copies of earlier pairs of the kind. And their stability over time is due to facts about their shared genetics, environments, physiological facts about how organisms maintain themselves over time, and so on. In other words, the uniformity and stability of dogs as a kind is grounded; there are mechanisms and laws which explain why they possess the properties they do and why those properties co-occur across members.

Four clarifications about the view are in order before continuing on. First, while the view speaks of natural kinds, this should not be thought to be drawing a contrast between naturally occurring and artificial kinds. Human made artificial kinds, like hammers and dollar bills, can be just as natural (in the sense relevant here) as those that are naturally-occurring, like electrons and asteroids. For example, the MacBook on which I am writing this dissertation is a member of a natural kind on the epistemic theory. MacBooks as a kind are both highly uniform and stable owing to the fact that MacBooks are all produced at the same factories according to the same design, and that later designs are partial copies of earlier designs. Due to this, there is a good deal of inductions you can draw regarding the size, shape, materials, and functions of MacBooks. Due to this, some proponents of the epistemic view eschew the term “natural kind” entirely for alternatives like “real kind” that do not carry the misleading connotation that natural kinds must be naturally occurring kinds.

Second is that whether something is or is not a natural kind is a matter of degree on the epistemic view. Further, there may be marginal cases where it is indeterminate whether a kind is natural or not. This is because the naturalness of a kind is a function of that kind's inductive potential, and inductive potential itself comes in degrees. The kind *planet* prior to Pluto's demotion was clearly a natural kind to some degree: it supported a great many inductions and formed a core part of various successful scientific theories over several centuries. However, with the discovery of other objects in the Kuiper Belt that counted as planets but bore little resemblance to the only known members of the kind up to then, we recognized (or at least should have recognized) that the kind was less uniform, and thus afforded less inductive potential, than we had expected.²⁷² By dividing the previous kind *planet* into planets proper and dwarf planets, we shifted our focus to kinds which afforded more inductions and thus, hopefully, would be more fruitful for scientific purposes. While the previous kind *planet* was indeed natural, it was not quite as natural (or not quite as real, if one would like) as the new kind of the same name.

Third, and perhaps somewhat oddly, whether and to what degree a kind counts as natural may be relative to the cognitive and perceptual capacities of the one drawing the relevant generalizations. Non-human animals plausibly cannot represent a wide variety of kinds in thought, e.g. quarks and Quakers, and thus cannot draw any inductive generalizations regarding them. In some cases, they may not even be able to represent rather mundane kinds in perception, and thus will likely lack corresponding cognitive representations as well, e.g.

²⁷² The reality of the IAU's decision to demote Pluto is, from what I understand, messier and more political than this ideal. At least part of the reason was worries about the number of planets schoolchildren would need to learn if the Kuiper Belt objects were included. The general idea should, however, be clear from the example.

cats and red objects as a kind. Humans may well be in a similar position, though with kinds we, for whatever reason, may not be in a position to fruitfully investigate. That is not to say that these kinds do not exist, or only exist “for” these sorts of thinkers, whatever such a phrase may mean. But what can be said is that they lack salience for these thinkers; they do not factor in any of their practical activities (including scientific and lay theory-making), and they are irrelevant to explanations of their thinking and behavior. They are, in a way, mere idle curiosities for creatures such as us, if we can even entertain them in the first place.

Finally, whether and to what degree a kind is natural can change over time; it is not an immutable property of the kind. The uniformity and stability of a kind can increase over time. Earlier models of cellphone differed radically in their designs, but consumer tastes have, over time, caused the kind to become far more uniform. Designer dog breeds may be another example, where the diversity of a breed is lost over time as breeders pursue a small set of traits taken to be representative of the breed. By contrast, a kind can decrease in uniformity and stability as well. Dog breeders may start to introduce greater genetic diversity out of a moral consideration for the health and wellbeing of the dogs they breed, and consumers may tire of standardized cellphone designs and demand greater variety in options.

A full defense of the epistemic view of natural kinds is beyond the scope of this chapter.²⁷³ I will, however, say a few brief words in its defense.

Firstly, one major desideratum for any theory of natural kinds is that it count as natural the kinds referred to in the statements of well-established scientific theories. If one’s theory of natural kinds fails to count, for instance, invertebrates as such a kind, that is effectively a

²⁷³ Though see Khalidi (2013, 2023); Boyd (1999); Kornblith (1993); Machery (2009), pg. 230-239; and Millikan (2017), pg. 11-26, for defenses of the position.

reductio on the theory. If any kinds are to count as natural, it is going to be those of the sciences, the kinds that afford our most secure and systematic knowledge of the world.

By tying together a kind's status as natural to its inductive potential, the epistemic view requires that a kind, by definition, be the sort which can factor in a successful science. It is only the kinds about which a great deal can be learned that science should focus its attention upon. And it is only those kinds with a great deal of inductive potential which can form the basis of the sorts of explanations that scientific theories traffic in. Thus, the proponent of the epistemic theory need not worry that the kinds referred to by successful scientific theories will be left out; by definition, they will not be excluded.²⁷⁴

Second, it does not much matter for my purposes if the epistemic view fails to capture what many metaphysicians have wanted out of a theory of natural kinds. One way to see this is to look at more traditional views and why supporters of them may balk at the epistemic theory. Natural kinds, many traditionalists have thought, should correspond to the "objective joints in nature," however that may be understood. But it is far from clear that the kinds with high inductive potential correspond to these objective, metaphysically distinctive kinds. In other words, it is far from clear that nature is carved up along inductive lines, or that the joints in nature form the boundaries between uniform and stable clusters of particulars bound together by underlying laws and mechanisms.

²⁷⁴ Of course, there may be kinds that the statements of scientific theories refer to that lack sufficient inductive potential. Phlogiston or the luminiferous aether presumably afforded few inductions, contrary to the hopes of those that posited them. But this is less a problem for the epistemic theory of natural kinds than it is an indictment of those scientific theories and their focus on the wrong, e.g. non-natural, kinds.

Worse, there is at least one reason to think that the kinds with high inductive potential do not correspond to traditional natural kinds. While the idea of a joint of nature is, to put it mildly, obscure, at least one plausible condition on them is that whether something is a joint in nature is not perspective dependent. A natural kind on this way of thinking cannot be a natural kind for me but not for you, or natural for us but not for our martian counterparts. It must be natural in some way that is wholly independent of whatever we or anyone else may think of it.

Natural kinds on the epistemic view, however, lack this feature. This is because, as earlier mentioned, the inductive potential afforded by a kind will depend in part on the capacities of those engaging with it. There are many inductive generalizations we could draw were we not limited by our working memory and reasoning abilities, the perceptual capacities available to us, and features of our epistemic position. A colorblind person or species lacks the capacity to draw inductions regarding the colors of some of the objects around them (at least without technological assistance), and a person with achromatopsia lacks this capacity regarding all the objects around them.²⁷⁵ And there is a good deal we likely could have learned about many kinds now lost to history, such as the supposed dialogues of Aristotle, or the lives and societies of many countless prehistoric peoples. But time has permanently washed away any of the basis we could have had for the relevant inductions.²⁷⁶

²⁷⁵ Again, as with Hellen Keller, this is not entirely true. Given a supportive social environment and people to inform them of a thing's color, they can acquire the relevant knowledge (if not knowledge of what it is like to experience it).

²⁷⁶ For instance, it may have been true (as some commentators alleged) that Aristotle's dialogues were superior to Plato's in terms of the quality of their prose, or that they bear certain commonalities in style and substance. Without the works themselves, however, we cannot perform the relevant inductions and must rely instead on mere testimony.

I do not, however, think that this failure of the epistemic theory to capture what more traditional metaphysicians have wanted is a problem. My goal here is to articulate a taxonomy and theory of conceptual representations that can do the sort of explanatory work expected of concepts by both psychologists and philosophers. And trying to come to a taxonomy of representations that includes kinds which maximize inductive potential is one way of doing this. My project here is largely guided by pragmatic and explanatory concerns, and the epistemic theory is primarily focused upon just these sorts of concerns, rather than the concerns of traditional metaphysicians.

Returning to concepts, in the normal case concept kinds individuated in terms of fine structure will fail to constitute natural kinds. This is because in the normal case concepts display little uniformity or stability in their fine structures, i.e. concepts are highly variable both between and across people, as I have repeatedly argued. What our concepts represent shifts constantly as we learn new things about the world and adjust how we think in light of our experiences. And each of our concepts will reflect our own idiosyncratic experiences, needs, and ends. Many concept kinds at the level of fine structure will have very few instances, and there will be little one can say about these kinds beyond the properties used to pick them out. A concept type like *prototype of birds that weights having a beak at 0.7, flying at 0.2, and eating worms at 0.1* is not likely to be widely possessed, and there is unlikely to be much of anything interesting that one can say about that concept that one cannot already say about bird concepts and bird prototypes more generally. Such a bird prototype is less like the natural kind *MacBook* and more like the non-natural kind *MacBook with a pink skin on its lid and a missing 'l' key*. It is not a concept kind which will afford much in the way of inductive potential.

5.2.3 – Natural Fine Structural Kinds

This is by no means true of all fine structural concept kinds, however; there are examples of conceptual representations individuated in terms of their fine structures that nevertheless may constitute natural kinds.²⁷⁷ Developmental psychology has found that the folk theories of humans represent the same domain specific-principles, e.g. humans represent artifacts as having their functions essentially, while animals are held to have their ancestry essentially or possess some deep, hidden essence.²⁷⁸ Further, there are empirical studies like Barbara Malt's study of the folk concept of water that show that at least some concepts do have the kind of stable, grounded uniformity characteristic of natural kinds.²⁷⁹

There will be at least three cases where concepts individuated in terms of their fine structures will constitute a natural kind:

1. Acquiring the concept is only possible in a limited number of ways, owing to either the nature of what is represented by it or the epistemic position of potential possessors.
2. The concept is typically acquired by being taught, and the method and manner by which it is taught is highly uniform and stable.
3. The concept is innate or is acquired as a result of processes or mechanisms that are themselves innate.

²⁷⁷ That there is any possible theory of what I am calling fine conceptual structure, and thus the exercises of the capacities they support, has been denied by some philosophers. See Rey (2009) for one example, though Rey may be merely denying that such representational structure is conceptual in nature.

²⁷⁸ Bloom (1996), Gellman & Wellman (1991).

²⁷⁹ Malt (1994). Though it is important to not assume such a concept is widely possessed, given that the subjects were all from the United States and of college age. It is likely that even concepts that are shared by smaller groups of people such as this still afford a good deal of inductive potential, albeit less than if such concepts were more widely shared across age and cultural groups.

Examples of (1) appeared earlier in the chapter. Most people likely have concepts of Aristotle that are near identical in terms of their fine structures, simply due to how our methods of knowing about him limit the kinds of properties our concepts can represent him as plausibly having. For (2), the common concept of Einstein which represents him as being a physicist who invented the atomic bomb is uniform and stable owing to the fact that people learn it from others in their community who know little of Einstein. And the domain-specific principles discussed by theory-theorists are plausible examples of (3), given both their ubiquity and how each is acquired at similar stages of development across all humans.

There are also likely some generalizations that can be made about conceptual representations, at least across the members of a single species. For instance, human vision is the perceptual system that provides the most informationally rich and reliably accurate representations to us. Thus, it would not be surprising if visually available properties are over-represented in human conceptual representations in a way they are not in the conceptual representation of species for whom vision is less central to their form of life.²⁸⁰

Other generalizations may be possible due to some concepts being constrained by how few relevant or distinguishing properties their represented kinds have. Concepts of sugar are highly likely to have sugar's sweetness as one of the properties represented. In part this is because it is the property of sugar that is usually most salient and factors in human desires, but also because it is the primary distinguishing characteristic of sugar versus another common kitchen staple, namely salt. It is not a coincidence that in the absence of clear labeling or

²⁸⁰ Dogs, for instance, utilize olfaction to a much higher degree than humans, and while they may be unable to recognize themselves via vision (regularly failing the mirror test for self-recognition), they may be able to recognize themselves via their distinctive scent. See Horowitz (2017).

packaging, tasting the substance is the best test for categorizing sugar and salt respectively as such. Further, even if there is widespread variation in fine conceptual structure, there may be some features of them that are widely shared, e.g. CLEAR and LIQUID are likely to be common elements of prototypes of water because of their high degree of availability for us, though our water prototype's fine structures may differ in other respects.²⁸¹

Thus the point is not that the fine structures of concepts will never be of explanatory value, that concept kinds defined by such structures offer no inductive potential, nor that respective concept kinds have no place in one's scientific theory. The point is that the explanatory value of such kinds is not a given, that natural concept kinds defined by their fine structures are often the exception, rather than the rule. The claim of high-level atomism that concepts are individuated in terms of their fine structures needs to then be amended ever so slightly:

High-Level Atomism 2: Concepts are typed referentially but have multiple subtypes individuated by their coarse structures, and *in some cases* these subtypes in turn have subtypes individuated in terms of their fine structures.

These considerations apply *mutatis mutandis* to the conceptual structures of my own view: in some cases, uniformity in the fine structures of various conceptual structures will result in them constituting a natural kind, the sort of kind that may be of explanatory value.

²⁸¹ Though even in cases such as these, one should be careful not to over-generalize. Hellen Keller likely represents water primarily in terms of its feel, given the unavailability of properties like clear and liquid. The tendency to universalize the cognitive content of concepts ought to be resisted whenever possible, always being guided by explanatory concerns of some kind rather than projection from one's own ways of thinking.

This may seem like a fairly minor change. But the fine structures of concepts have long been expected to explain a great deal. And in some cases, this supposed explanatory value has been used to argue backwards to the uniformity of such concepts, e.g. that the existence of normative demands on concept possessors necessitates a common fine structure that guides and coordinates the activities of competent concept possessors. Recognizing both the widespread variability of concepts and the often limited explanatory value of their fine structures thus serves to ward off the temptation to think that concepts can do more for us and our theorizing than they in fact can. This theme continues in the next argument.

5.2.4 – The Functional Role of Concepts

One curious thing about the debate between philosophers and psychologists on the matter of concepts is that the concepts posited by each side of the debate often bear little resemblance to one another. This is so even at the high levels of abstraction that these debates tend to occur at. The concepts posited by philosophers are, ontologically speaking, quite variable, ranging from being mere abstracta (such as Fregean senses) to abilities, along with the mental representations more familiar to psychology. And these concepts sometimes have features, such as their assumed accurateness, and play roles, such as acting as building blocks for thought, capturing some privileged way of thinking about the concept's referent, and serving as the basis for philosophical analysis, that are entirely alien to psychology.

Psychologist's concepts, by contrast, are in nearly all cases structured mental representations.²⁸² The ontological variability present in philosophy is thus largely absent from psychology. Further, these concepts play quite different roles in the theories of psychologists

²⁸² This understanding of the ontology of concepts is not, however, universal in psychology. See Rosch (1978, 1999) for one such dissenting position.

than those of philosophers, with this difference in explanatory role possibly explaining the greater agreement among psychologists when it comes to the ontological category concepts fall within. The concepts of psychologists act primarily as stores of information about the world around us and function to make that information available for use by the processes underlying an array of central, practical activities that many (if not all) thinking creatures engage in. And far from making available to the thinker some universal, privileged (and often true) perspective on its referent, psychologists' concepts characterize the idiosyncratic (and often mistaken) perspectives of their possessors. Conceptual variability, inaccuracy, and even inadequacy, are the norm, rather than the exception, on psychologist's theories.²⁸³

It is not clear what to make of this dissimilarity in what concepts are held to be by both sides of the debate. One possible reaction is that this disparity arises from each side focusing on different aspects of a single phenomenon. On this understanding, concepts both function as thought constituents and as stores of information to be exploited by our cognitive processes. It is simply that, up until now, philosophers have focused only on one of these two theoretical roles and psychologists the other. If this is so, then what is needed now is some way of reconciling these two understandings. This would involve figuring out how a single mental representation can be both simultaneously, and what the broader cognitive architecture containing such representations must, or at least does, look like.

Another possible reaction is to hold that this shows that philosophers and psychologists are talking past each other, with each side focused on different, though perhaps related,

²⁸³ Psychologists may not, however, see it in quite these terms. Often what seem to philosophers to be mistaken category judgments or false beliefs about a world-aspect are to psychologists simply evidence of a difference in concept rather than a defect in concept. I will discuss this way of thinking about concepts in ch. 6 and some issues it may give rise to.

phenomena. Perhaps such things as prototypes and exemplars are mere conceptions, rather than concepts proper, with conceptions capturing idiosyncratic perspectives and concepts some privileged one.²⁸⁴ Or perhaps they are representations different in kind from those including thoughts and their constituents, perhaps forming part of an entirely separate cognitive system. It may even be that concepts as constituents of thought can be dispensed with entirely.²⁸⁵ Perhaps, for instance, prototypes and the like are used primarily in fast, automatic, and unconscious processes, while the background knowledge that concepts proper constitute (concepts in the philosopher's sense) is used for slower, more deliberate, often conscious reasoning.

The right reaction is, I think, somewhere in the middle of these two. On my view, the conceptual atoms of philosophers are not the same thing as the conceptual structures of psychologists. They are two distinct kinds of mental representation, differing in both form and function. The second reaction is then, at least on this point, correct: psychologists and philosophers are not theorizing about the same things, and they are expecting incompatible things from concepts.

But those philosophers who take this path often oversell the difference. Conceptual atoms and structures are not wholly separable, consigned, say, to different cognitive systems, processes, or even different ontological realms. They are symbiotic and mutually dependent on one another. Mere conceptual atoms would be, as has long been argued, worthless to their possessors. And conceptual structures could not even exist in the first place without the basic elements from which they are constructed. A conceptual atom without a structure

²⁸⁴ Rey (1983, 1985).

²⁸⁵ Machery (2009), pg. 26-28.

encompassing it is akin to a generator without wires to transfer the power generated, or a heart without arteries and veins to transport blood.

The first reaction then is right in seeing philosophers and psychologists as each taking up differing perspectives on one and the same broad phenomenon. It is not a single kind of representation they have each directed their gaze upon, but rather different aspects of a single complex and many-faceted cognitive system (or systems). It is a system(s) consisting of simple, unstructured representations of the many varied aspects of the world, as well as a multitude of structured representations formed from these to meet the varied needs of the thinker.

Additionally it is not clear that there is any explanatory need to posit beliefs or PA-supporting representations more generally as something over and above or distinct from representations like prototypes and theories. The various representations posited by psychologists are capable of representing a wide variety of phenomena in a wide variety of ways. Definitions can represent the properties had necessarily by members of a kind as well as those sufficient to belong to it. Prototypes can represent the general tendencies of kinds, such as through the properties typically had by the kind's members. Exemplars can represent particulars and their properties both across and at particular times. Theories can represent causal, functional, and other such relations between phenomena as well as their domain membership and the domain-specific principles governing them. Ideals can represent the thinker's preferences as well as what they take to be the best or most desirable instances of a kind (whether real or imagined).

While there are of course beliefs whose content is not plausibly captured by any of the above, the class of concepts is open-ended for a reason. Anything not captured above can be captured by as yet undiscovered forms of conceptual representation. And this is hardly new.

The shortcomings of definitions brought about the invention of prototypes and exemplars, the deficiencies of prototypes necessitated the introduction of exemplars, and the shortcoming of both prototypes and exemplars led to the positing of mental theories. The supplementation of concepts by psychologists with ever more coarsely structured concept kinds has been the norm for decades, and there is little reason to suppose more will not be posited in the future. Through such expansions, the range of belief contents capturable expands in turn.

5.2.5 – Some Objections

It may be objected that while conceptual structures could be plausibly held to perform the same functions as beliefs, there is nothing that looks to be playing the role of other propositional attitudes such as desires. Concepts in the philosopher's sense are supposed to be, not just the constituents of those representations supporting our beliefs, but the constituents of those representations supporting all of our various PAs. Prototypes, exemplars, and the like all appear to be broadly *descriptive* representations; they represent how the world is taken to be by their possessor. There appears to be nothing corresponding to *directive* representations, such as whatever underlies desires, wishes, and hopes, among other PAs.

While this is true, this need not be a problem for my view. While conceptual structures are defined as psychologists define concepts, i.e. as representations factoring in processes underlying core cognitive capacities, what exactly counts as a core cognitive capacity is left largely open. There is nothing prohibiting including capacities relevant to motivation and preference and thus countenancing conceptual structures that are directive rather than descriptive in nature. It may even be the case (though I will not argue the possibility here) that some of the familiar conceptual structures from psychology could, with a change in functional

role, serve as directive conceptual structures as well.²⁸⁶ Some, such as ideals, already represent the evaluations and preferences of their possessors, and with some tweaks could plausibly function to not just represent one's preferences, but to motivate one to satisfy such preferences as well.

Another concern may be that the functional role of concepts in psychology is being misconstrued here. While it is uncontroversial among psychologists that concepts factor in processes like categorization, it is sometimes claimed that they play a much more limited role than I have been ascribing to them. Take for instance Machery's definition of concepts, which I have relied on from the outset save for two seemingly small omissions²⁸⁷:

I propose that in psychology, concepts are characterized as being those bodies of knowledge that are *stored in long-term memory* and that are *used by default* in the processes underlying most, if not all, higher cognitive competences when these processes result in judgments about the referents of these concepts.²⁸⁸

The first of these omissions is the requirement that concepts (or in my case conceptual structures) be stored in long-term memory. This is to accommodate the possibility of what

²⁸⁶ This is largely in agreement with standard functionalist views of mental states, where the difference between beliefs and desires with the same cognitive content is due to their respective functions, with the propositions they express being common across both.

²⁸⁷ There are two other, more minor, changes. First is that I have throughout referred to core cognitive capacities as opposed to competencies. This is not, I take it, a substantial philosophical difference (or at least it was not meant to be). Rather, it is more out of worries regarding the normative connotations that "competency" carries, and a desire to avoid such connotations whenever possible. Second is that I have chosen to speak of "representations" rather than "knowledge." This is because I want to avoid the implication that concepts are necessarily accurate. Importantly Machery does not intend this implication either, but rather uses "knowledge" in the way many psychologists do, such that one can know P even if P is false.

²⁸⁸ Machery (2009), pg. 4.

psychologists call ad hoc concepts.²⁸⁹ These are concepts that are assembled on the fly in short-term memory for a particular task or purpose and then discarded when no longer needed. I may, for instance, assemble an exemplar of a piece of clothing at the store I am undecided on buying, along with a cognitive map marking where in the store I left it. Or I may assemble an ideal during the process of trying to determine where to go out for dinner, representing what would for that night be ideal and then discarding said representation when the night is done. There is, so far as I can tell, little compelling reason to exclude such short-lived concepts from the category altogether, especially given both their structural similarity to representations stored in long-term memory (a similarity plausibly due to identity) and their seeming ubiquity. But while of independent interest, it is not this omission which is of greatest relevance here.

A more pressing omission is that concepts are those representations “used by default” in the cognitive processes underlying core cognitive capacities. While what counts as “default” is somewhat murky, Machery’s characterization is clear enough for our purposes. Speaking of knowledge stored in concepts that is to be relied upon by default, Machery says:

The knowledge that is stored in a concept of x is preferentially available when we think, reason, and so on, about x. So to speak, it spontaneously comes to mind. By contrast, the knowledge about x that is not stored in a concept of x is less available—it does not spontaneously come to mind. The knowledge that is not stored in a concept of x is used only when the knowledge that is stored in this concept is insufficient or inadequate for the task at hand.²⁹⁰

²⁸⁹ I discuss ad hoc concepts in more detail in § 2.4.2.

²⁹⁰ Machery (2009), pg. 11-12.

So on Machery's view, concepts are limited to those representations that are drawn upon first in the performance of some task, while the remainder of our beliefs are drawn upon only when our concepts have proven not to be up to that task. This, in effect, is to limit concepts to providing explanations of processes and capacities such as those involved in making snap judgments, i.e. quick, largely unreflective judgments made usually due to the presence of tight time pressures, or the relatively low importance of the matter being judged. Background beliefs are what explain our slower, more reflective, reasoned judgments, those judgments being made when we have more ample time to consider matters or when the stakes are high enough that relying on a snap judgment would be an unacceptable risk. The view is then committed to a robust distinction between conceptual and background knowledge.²⁹¹

If conceptual structures are nothing more than concepts as understood by Machery, then they cannot play the same functional role that beliefs in general can. The two would factor in different processes, and the beliefs supported by concepts would be only a subset of one's total beliefs, with the remainder forming the background. There are, however, two reasons why I do not think this additional condition Machery places on concepts is advisable. As such, there is I think no reason to believe that conceptual structures are quite so limited in their functions, and that they can play the role of PA-supporting mental representations.

First is that while Machery's characterization is a natural fit for some types of concepts it seems far less plausible for others. For some conceptual representations, like theories, it does not seem to fit well at all. In many of the experiments carried out by theory-theorists, subjects are given ample time to reason about and reflect on the situations presented to them before

²⁹¹ This distinction and its issues are discussed in § 3.3.1.

making their judgments. Similarly, definitions seem much more likely to be drawn upon to supplement other representations than the reverse, and thus to form part of one's background knowledge rather than one's concepts on a view such as this. On many theories in psychology, such as hybrid and dual process theories, definitions play this role explicitly, forming the cores of concepts which are only exploited when the concept's periphery (usually a probability-based representation such as a prototype or exemplar) has proven inadequate in some way. But to the extent they do play such a role, they are not being utilized by default and are thus not concepts, contrary to the statements of the psychologists who posit and theorize about them.

Those concepts that Machery's definition best captures are prototypes and exemplars, which do seem in many cases to be used for quick, largely unconscious reasoning. When you need to make a categorization judgment under strong time pressures, relying upon the typical properties of a kind is a natural and more often than not reliable method. If I find a shiny yellow metal in the bed of a stream, a judgment that it is gold based upon a prototype is both natural and reasonably likely to succeed. It is true that it may well be fool's gold, and that merely basing a judgment on such superficial properties is prone to failure in some, or even many, cases. It is also true that it would be far more accurate to rely instead on beliefs about gold's atomic number or beliefs concerning the chemical tests by which gold can be more reliably detected. But when knee deep in the water with little to lose were you to judge wrongly, you need speed and ease of use more than you need raw accuracy. The beliefs are only useful for the purposes of categorization insofar as the properties they represent are available to the thinker, and in this case and many others they are not. It is certainly plausible that in such a case a prototype is used by default in the way Machery suggests.

Why though should one suppose that this is always the case? Could it not be, for instance, that what counts as default for one person is different than for another? Perhaps the prospector draws upon a prototype of gold representing its typical, observable properties like shiny and yellow; after all, this would be most useful given their need to recognize gold on the basis of its visually accessible properties in order to prospect well. But could not a chemist rely by default on a definition of gold representing it as having the atomic number 79? Presumably such a concept would better serve the chemist, as in many of the tasks required of their profession, relying on gold's superficial properties even in the earliest steps of their work would be both inappropriate and a waste of the chemist's time. Or could they instead rely on something far simpler and more available, namely that samples of gold in the context of labs, unlike gold lying in streams, tend to be labeled as such. In such a context, beliefs about what gold is called and the labels on samples may seem more likely to be default. Why would someone retrieve by default bodies of knowledge that would be effectively useless within the context they find themselves in, only to then after immediately retrieve knowledge that *is* relevant to and useful in that context?²⁹²

This example thus serves to illustrate the second problem, which is that there is no non-arbitrary way to draw the line between which beliefs are drawn upon default and which are not. This is, in part, because what beliefs count as default will be a function both of the task

²⁹² Worse, such thinkers may also know that these bodies of knowledge are useless within that context. A chemist may know that the superficial properties of gold are sometimes misleading and take caution not to rely on them, and a prospector may know they are unable to use anything but such superficial properties (even if they know they know its deep properties), and thus have no choice but to rely on them. This is common and can happen in nearly any context for nearly any task. If an acquaintance has a twin, I ought to know not to rely too much on their superficial properties to make a judgment as to their identity.

being performed and the context in which that task is taking place.²⁹³ The chemist is likely to draw upon a definition of gold when at an academic conference and that definition is being utilized to comprehend the utterances of their peers; relying on a prototype in such a case is liable to lead to misunderstanding of what is being said. The prospector is much less likely to understand “gold” in this way, and if they are old-timey in addition then they will not even have the possibility of doing so owing to their lack of chemical knowledge.

It is this additional issue, that people may have not even overlap in their beliefs about some aspect of the world, that makes the issue especially problematic. I have not the foggiest idea what ruthenium looks like, but I know that it has the atomic number 44. In fact it is all I know of the metal.²⁹⁴ What else but the definition could be drawn upon by default in my case if it is all I know of it? I cannot draw upon by default a representation I do not possess. And I could have instead learned of ruthenium by seeing a sample of it in a museum and forming a prototype or exemplar of it, perhaps not learning of its atomic number due to never reading its accompanying placard. In such a case, the prototype would be utilized by default instead simply because nothing else could be. So depending on what I have or have not learned, either I will draw upon by default a prototype of ruthenium or a definition of it. These points all generalize. What will be drawn upon by default depends on everything from what you are trying to do to what you know, and thus what counts as default will vary according to person, time, goal, and various other widely varying factors.

²⁹³ Barsalou (1993); Sperber & Wilson (1998); Roth & Shoben (1983).

²⁹⁴ Excepting of course that it is a metal, that it is a chemical substance, that it takes up space, and other such truths entailed merely by its being a metal.

Thus, on this view a prototype could count as a concept in your head but not in mine. And a definition can be a concept in one moment when I am engaged in a particular task and cease to be at some later moment when I switch to do something else for which it is less immediately suitable. Concepts so understood strike me as strange, chimerical entities, ones which I cannot see the utility of positing. What seems more plausible is that concepts are simply whatever representations are drawn upon by core cognitive processes, whether by default or not, and which concept gets drawn upon first is a matter of what is available in the world via perception and what conceptual information is appropriate to the context and task at hand. A story would need to be told of how it is determined which concept is utilized when and for which tasks, but there seems no reason in principle such a story could not be told.²⁹⁵ To restrict concepts to default representations seems to bring with it little in the way of explanatory gains and to run contrary to the practice of many psychologists, all while incurring the cost of needlessly complicating the typology for concepts and other potential conceptual representations.

An additional, perhaps even more serious, worry is that the very notion of a default belief may itself be questionable. The definition implies that there is some sharp cutoff in cognitive processing when the initial beliefs drawn upon (the concept) have proved wanting and now new beliefs are retrieved from long-term memory (one's background beliefs). When these background beliefs are drawn upon, new cognitive processes corresponding to them are engaged. So there is a discontinuity, both in the representations being used and the processes

²⁹⁵ I will attempt to give the beginnings of such a story, one inspired in large part by Millikan (2017), in ch. 7.

that make use of them, and this discontinuity marks the boundary between those beliefs that are properly conceptual and the remainder of one's beliefs in the background.

This does not, however, appear to be what actually occurs. Rather, there is in many cases a smooth and steady increase in the beliefs brought to bear on the task the longer one is engaged in it.²⁹⁶ Consider for example a simple task like finding a particular brand of shampoo at the supermarket. At first you rely on the simplest and most easily discerned properties, such as the color and shape of the bottle. These properties you can see at a far distance, and they will be the first available to you in making a categorization judgment. If this fails (perhaps because the manufacturer switched bottle designs recently), you will have to draw upon something else, such as your knowledge of the brand and product names, looking for these on the various bottles in front of you. Of course, even this could fail (perhaps the name changed along with the bottle design). In such a case you may be able to find and correctly categorize the shampoo by looking at the ingredients listed on the back of the bottle, relying on deeper properties concerning the shampoo's formula, rather than superficial properties concerning its container. Though even these properties may lead one astray.

Perhaps there are still discontinuities between the different identifiable strategies being utilized, such as that between utilization of a prototype and similarity-matching process to a definition and a rule-checking process. But even so, there is here not one discontinuity but many. It looks, I suggest, less like a shift from relying on concepts to background knowledge, and more like shifts between different conceptual structures and the strategies they make possible. Further, the order in which these strategies are utilized need not correspond to what

²⁹⁶ See § 3.3.1 for more discussion of this issue.

Machery argues for. In some cases and for some tasks one may rely first on a prototype and then a definition, while in others it may be the definition first utilized with the prototype available as backup.

5.3– Conceptual Atoms and Conceptual Structures

I have been arguing in this chapter that the concepts of both philosophers and psychologists do not seem to be functioning and doing the sort of explanatory work one would expect were they one and the same. The prototypes, exemplars, theories, and the like of psychologists appear to be doing what the PA-supporting representations composed of concepts (as understood by philosophers) are doing, rather than acting as one would expect of the components themselves. One, I think, plausible explanation for this fact that I will pursue here is that this similarity in function between PA-supporting representations and the concepts of psychology is due to their identity. Philosophers and psychologists have not, as many seem to have thought, been engaged in a debate over the components of thought. Rather, in their theorizing philosophers have latched onto the components of thought and psychologists to the composed thoughts themselves.

If so, this reveals a problem: psychologists and philosophers simply are not talking about the same things when they use “concept.” And if this is so, it may be of use to start angling for some terminology and a new taxonomy of the conceptual domain to go with it. I propose to attempt such a thing in what follow. On the view I will develop here, concepts divide into two distinct kinds of representation: *conceptual atoms* and *conceptual structures*, each of which performs distinct, though related, functions. It is my hope that bifurcating concepts in this manner can simplify much of the theorizing surrounding concepts and serve to ward off some of the more puzzling problems more traditional views have generated.

Conceptual atoms are atoms in the classical sense, as they appear in earlier unattenuated forms of atomism. They are referentially individuated and semantically unstructured mental representations which function to combine into complex conceptual structures. These structures form a broad and open-ended class of structurally individuated conceptual representations that includes prototypes, exemplars, theories, and whatever other representations psychologists have or will someday need to posit to explain the various core cognitive capacities and their exercise. In so combining atoms form relations between themselves which encode useful information about the world and make it available to be exploited by the thinker's mental processes.

I will understand each of these types of representation throughout this chapter and the next two as follows:

Conceptual Atoms: Referentially individuated and semantically unstructured mental representations whose function is to combine together with one another to form conceptual structures.

Conceptual Structures: Structurally individuated and semantically complex mental representations composed of conceptual atoms and which function to both (1) guide the exercise of core cognitive capacities through the information they encode in the relations between their constituent CAs and (2) to support the having of propositional attitudes by their possessors.

Figure 5.1 shows the relationships between the various types of conceptual representation on my view. As may be seen, it is in some ways more an elaboration of high-level atomism than a replacement for it, positing another kind of conceptual representation to

serve as the constituents of what traditional atomism called “concepts.” In parentheses for each representation are the individuating properties for representations of that type and on that level of abstraction. That there are two CAs (that of x and either an atom of y or z) is meant to show that CAs do not occur individually within CAs but always as groups, as it is only by standing in relations to one another that CAs are able to encode exploitable information about the world. The distinction between core and peripheral CAs indicated in brackets will be touched on in a few pages, but for now all that needs to be known is that it is the core, and not the peripheral, CA that determines the referential content for the CS as a whole.

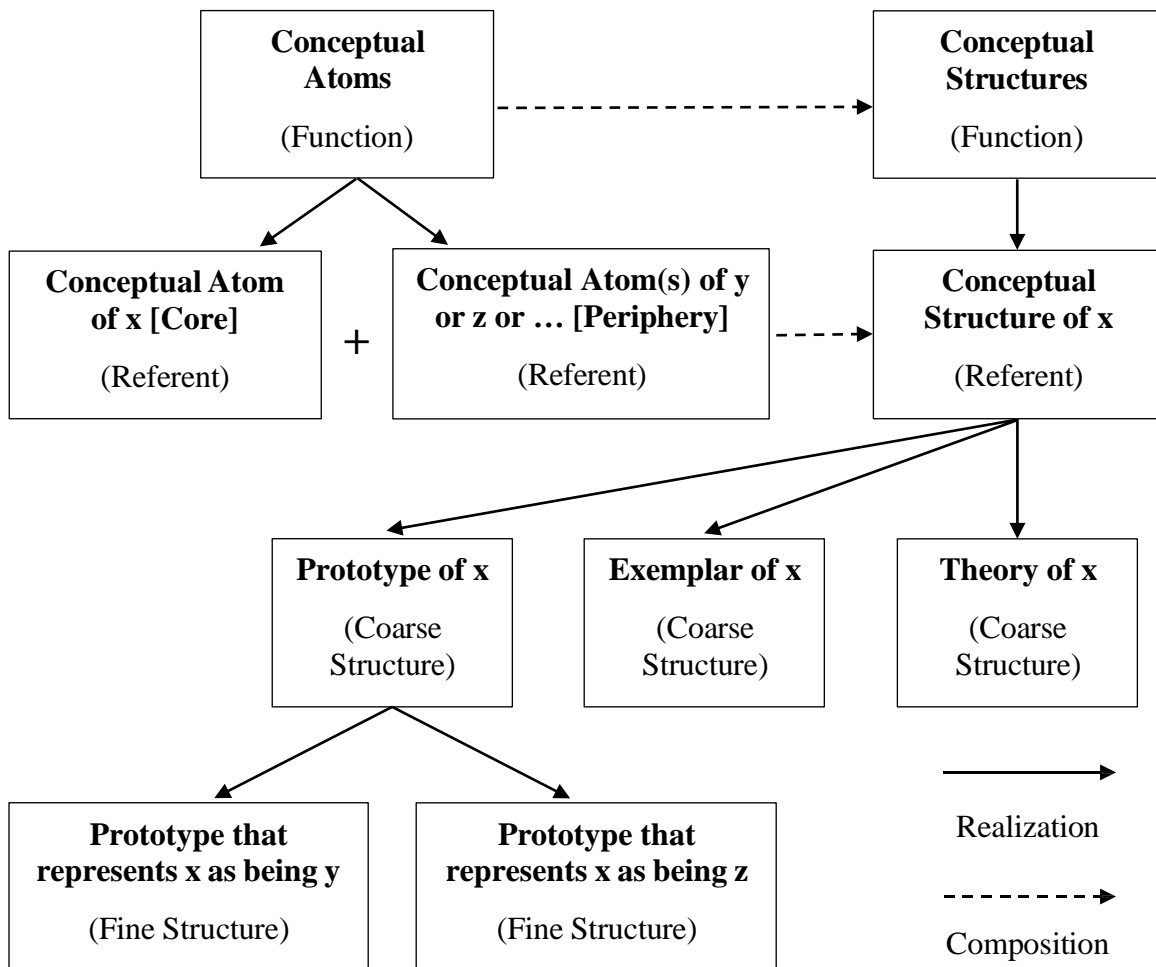


Figure 5.1

Consider a simple example: predator detection in domestic chickens. Male domestic chickens produce alarm calls upon identification of potential airborne predators, such as various raptor species, as well as engage in other forms of relevant behavior, e.g. crouching and visual fixation on the perceived threat. In one experiment, domestic chickens were placed in an environment with an overhead video screen that displayed the outlines had by common chicken-predators along with the outlines of various non-salient shapes. What was found was that the chickens responded with calls and predator-directed behavior upon perceiving the predator-shaped outlines, but not for the others. Further, the rate and magnitude of the chicken's warning calls, and their other predator-directed behavior increased with the apparent size and rate of movement of the predator-shaped outlines on the overhead screen.²⁹⁷

The chicken relies on the identification of highly superficial properties in order to identify the presence of some of its aerial predators. It looks for shapes of a certain size in the sky above, moving within a certain range of speed. For the chicken, this is a reliable enough mechanism. In the chicken's natural environment, objects that share the outline of the predator shapes used in the experiment tend to actually be the natural predators of chickens. Further, both a predator's apparent size and apparent rate of movement relative to the chicken correlate with the altitude of the predator relative to the chicken, and this with the danger it poses to the chicken. Thus, it is not surprising that the chicken's behavior displays a corresponding increase in intensity: as apparent size and rate of movement increase, the likelihood that a predator (and thus death for the chicken and its flock) increase as each of these factors increase.

²⁹⁷ Evans, Macedonia, & Marler (1993).

One lesson to take from this, one that has arisen multiple times before now, is that CSs do not need to be especially robust or complex in order to be adequate to the needs of their possessors. That the chicken is easily fooled by something as simple as a shape displayed on an overhead video screen says, I believe, little on its own about the chicken's cognitive sophistication, particularly in regard to what sorts of world-aspects the chicken can represent. The chicken lives a comparatively simple form of life, one which has as one of its major components the avoidance of predators. Its aerial predators tend to be relatively uniform, displaying similar properties (shape, size, speed) across encounters. And the manner in which the chicken engages with its predators is similarly simple. The chicken does not need to ponder the nature of its predators or their intentions. It merely needs to flee and warn the other members of its flock. A CS focusing on contingent but perceptually available and easily identified properties is, given the chicken's environment, needs, and perceptual capacities, highly effective (or at least, effective enough to keep enough chickens alive such that the CS persists over evolutionary time).

Such cases do not show the chicken is not *really* representing predators (perhaps only representing shape, size, and speed, i.e. the properties by which the identification is made) any more than the similar failures of human identification on Twin-Earth show that humans cannot *really* think about water. If chickens could not represent predators and toads prey by virtue of their respective tendencies to be deceived in experimental settings, it is unclear why humans, vulnerable in the same way (though perhaps requiring greater cleverness on the part of the architects of the experiment), would not be similarly representationally impoverished.

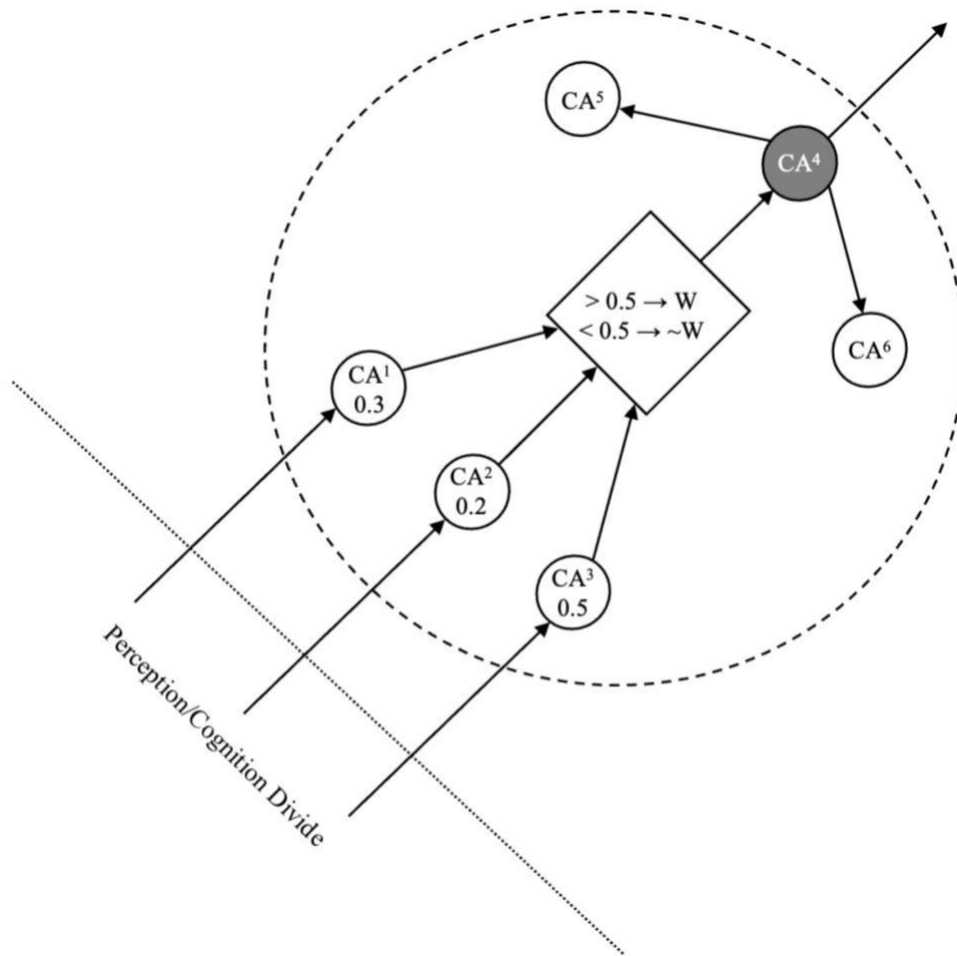
Consider now a fairly mundane human case where someone identifies a tool in a messy toolbox. Let us say that it is a wrench of a particular size being identified by means of a

prototype of wrenches. What makes something a wrench is its function, with this a matter either of the intentions of its designer or the effects wrenches produce that keeps them in production and thus ensures their survival as a kind.²⁹⁸ As with the chicken and the ancestral relations of the raptors circling overhead, the function of a wrench is not perceptually available, nor are the intentions of its designer and its effects on the wider consumer market.²⁹⁹ The identification in most cases will best proceed through the exploitation of contingent, perceptually available properties such as the wrench's shape, size, weight, color, and identifiable markings. That these properties tend to co-occur in the types of environments humans find themselves in provides relations between these properties that can be exploited by fitting representational strategies, such as the use of a prototype whose co-occurring features mirror the co-occurring properties in the wrench.

A simple model of such a prototype CS is given below:

²⁹⁸ Which of the two will depend on which theory of functions for human artifacts one endorses, either design or etiological functions. These theories among others will be considered in ch. 7.

²⁹⁹ One could, however, infer the likely function of a wrench via testing it for various uses and seeing which uses it is a good fit for. This recognitional strategy does, however, depend on the artifact actually being good at performing its function, and not performing two well at tasks for which it is not designed for. One could via such a strategy come to mistakenly believe that the function of a straightedge screwdriver is the opening of paint cans, rather than the driving of screws.



Suppose that CA¹ – CA⁶ have the following referential contents, weightings, and status as peripheral versus core:

CA¹ = WEIGHT, Weight 0.3, (Peripheral)

CA² = SILVER, Weight 0.2 (Peripheral)

CA³ = WRENCH-SHAPED, Weight 0.5 (Peripheral)

CA⁴ = WRENCH (Core)

CA⁵ = METAL (Peripheral)

CA⁶ = TIGHTENING AND LOOSENING (Peripheral)

The small circles each represent CAs, with the CA shaded in representing the core and the remainder representing the periphery. The larger circle with dashed outline represents the CS of which these CAs are constituents. The lines between the CAs and other components of the CS represent inferential, causal, or functional relations. The box in the middle of the CS represents the accumulator for the prototype, with the numbers representing the category threshold, the arrow a simple conditional, and the W and ~W representing positive and negative categorization judgments respectively. And finally the dotted line to the bottom left of the CS represents the division between perceptual and conceptual representation, with all falling to the left being perceptual and all to the right being conceptual.

CA¹ – CA³ represent the perceptually available properties which guide the retrieval of the WRENCH CS as a whole. These three CAs are functionally connected to perceptual representations which guide their retrieval, e.g. a perceptual representation of redness guides the retrieval of the CA RED. The accumulator is setup such that the sum of the weights of CA¹-CA³ must surpass the category threshold of 0.5 for CA⁴ to be retrieved. Each of these three CAs are weighted such that there needs to be the prior retrieval of, at minimum, the feature SHAPE and at least one of the two features COLOR and WEIGHT. When that threshold is triggered, CA⁴ is retrieved, the core CA for the CS, and a positive categorization judgment is made. Now with CA⁴ retrieved, the thinker can further retrieve CA⁵ or CA⁶, inferring that wrenches have properties that were not available to perception. And the core CA⁴ could then be itself used to guide the retrieval of some further CS (suppose, for instance, that CA⁴ is also a peripheral CA

for an exemplar of wrenches, with that exemplar itself forming a part of an exemplar set for tools more generally.³⁰⁰

A CS such as this could, of course, fail to be adequate given an abnormal environment: the manufacturer could have stamped incorrect markings, it may be a child's toy wrench instead of a real one, tricky psychologists could be conducting another experiment, and so on. And it may not be adequate given certain ends, e.g. it may be a socket wrench when what is called for by the task is a pipe wrench. But in normal cases, in normal environments, and given the typical needs of their possessors, the CS in this case is perfectly adequate. CSs do not need to be adequate across all possible times and places, real or otherwise. They merely need to be good enough in the here and now, for the purposes towards which we actually put them.

5.3.1 – Some Outstanding Issues

There remains, however, at least two dangling strands in this discussion. First is what determines the referential content for CSs, assuming that the referential content for their constituent CAs is known. On this matter, I will be brief, as what I say here is largely in line with more traditional views. I take it that the referential content of a CS is simply a matter of one of its CAs being embedded in the right way in that CS's structure. This is similar to how the sentence "Cats are adorable" is about cats because "cat" occupies the subject position of the sentence, or the belief CATS ARE ADORABLE being about cats because CATS bears the right syntactic or functional relation to ADORABLE. The cognitive content of a CS is determined in the same manner as that for a traditional concept, namely the relations between all of its

³⁰⁰ This is of course all greatly simplified. The processes intervening between CA¹ – CA³ and CA⁴ are likely much more complex than what is represented here, as this is based on some of the earliest form of prototypes posited. And there would be at least as much intervening processing between the inference from CA⁴ to CA⁵ or CA⁶.

constituents. Far more could be said on this issue. However, the actual structure of CSs is more properly investigated by psychologists rather than philosophers. And, as may have been noticed, what a given CS represents entirely depends on the referential content of its constituent CAs. If a CS represents *x* in virtue of a CA of *x* being properly situated in the CS's structure, then one cannot know what the CS represents absent a prior understanding of what the relevant CA represents. I will touch on this issue in chapters 6 and 7.

I will call whichever CA determines the referential content of a given CS, i.e. the CA that represents what the CS of which it is a part represents, the *core* CA relative to that CS. The other CAs that constitute the CS I will term *peripheral*. For instance, the CA DOG is the core CA for prototype CSs of dogs, while CAs such as FURRY, FOUR-LEGGED, and the like are peripheral CAs for such CSs. Further, whether a given CA is core or peripheral is relative to a given CS. Thus the CA DOG can be both the core CA for a prototype CS of dogs and simultaneously be a peripheral CA for a prototype of mammals. This could be for the same reasons “dog” can be subject in one sentence and object in another, i.e. one token of “dog” occupies the subject position of a sentence, and a distinct token of “dog” occupies the object position of a distinct sentence.

For the sake of simplicity, though, I will assume a network model for CSs in what follows. On this model, multiple CSs can share a single token CA, and CSs are perhaps better thought of as portions of a single CA network, rather than as discrete representations stored in memory.³⁰¹ The diagram below illustrates how this may work:

³⁰¹ It would, in this way, be highly similar to Millikan's (2017) uniceptual networks and unitrackers. As with what I am here assuming, Millikan sees unicepts (atomic representations) as forming a vast network and unitrackers consisting of portions of that network. Unitrackers

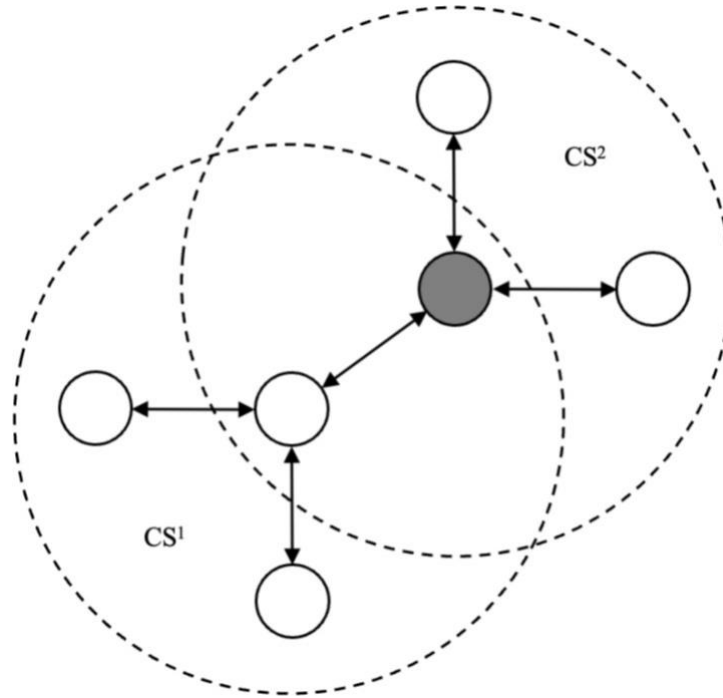


Figure 5.2

In the above diagram, CS¹ and CS² share two CAs in common, namely those CAs that fall within the overlap of the two. And the shaded in CA is a core CA for CS² while being merely a peripheral CA for CS¹. It is this sharing of CAs that I will argue in chapter 7 allows for the use of one CS to guide the retrieval and application of another, allowing for a more flexible picture of conceptual representations than one that requires they be deployed independently of one another.

The second dangling thread is the relation between CSs and the various propositional attitudes. Throughout I have talked of CSs and other complex conceptual representations “supporting” or “underlying” PAs. This language is, however, quite vague and unilluminating.

on her view are individuated by their proper function of identifying the referents of their corresponding unicepts and setting up/maintaining relations between that unicept and its peers.

As such, it would be good to have some more substantial account of just what this relation of “support” is. Unfortunately, this is not something I have, nor do I have the space to explore it here. At minimum, I think that for a CS to support a PA is for a thinker to stand in that PA in virtue of their possession of that CS. I suspect, however, that if more detail is desired, my own view could avail itself of the same answers that could be given for any other view that sees possession of complex mental representations as the basis for standing in propositional attitudes.³⁰²

5.4 – Concept Eliminativism

One worry, a common one in the literature, may be thought to ensnare my account here as well. Are what are called “concepts” just the same thing as what I have termed “conceptual structures” (or vice versa)? More pointedly, is this a mere terminological issue, a matter of mere theoretical bookkeeping? In a way, yes, insofar as it is somewhat arbitrary that I have chosen to eschew the use of “concept,” rather than assigning the word to either what I have termed conceptual atoms or conceptual structures. I could have followed the usage of psychologists and called prototypes, exemplars, theory components, and so on, “concepts,” while still calling the constituents of these representations “conceptual atoms,” or perhaps using a different term entirely. But I would then run into difficulties when conversing with philosophers, as what I would term “concepts” do not factor as constituents in the representations supporting our having of propositional attitudes, and thus are not what philosophers would be inclined to call “concepts.” Likewise, calling conceptual atoms just

³⁰² It is worth mentioning in brief that this is a more contentious issue than my passing discussion may suggest. It may be that, for instance, explanations of thought and behavior at the personal and subpersonal level are not systematically related. Or perhaps one’s PAs are a matter of the behavior that one’s CSs dispose one to engage in, rather than directly a matter of one’s possession of internal representations such as CSs.

“concepts” would similarly impede discussion with psychologists, who may be unclear how something that is incapable of anything that concepts do nevertheless qualifies as one. Given the sheer number of ways that “concept” is understood in the literature, to use the term “concept” is to invite confusion and misunderstanding, especially when one is trying to promote a more revisionary theory of conceptual representations.

I do not, however, think that this is nothing more than mere bookkeeping. What is substantive in the division I have drawn is the insistence that, whatever one calls “concepts,” those things cannot be both (1) constituents of PA-supporting representations, and (2) representations factoring in those processes underlying core cognitive capacities. There needs to here be a division: we need one type or representation for (1) and a distinct type of representation for (2). One and the same type of representation cannot pull double-duty, despite the long-held wishes of many philosophers and psychologists.

Perhaps use of these terms in the first place is already a mistake. Perhaps we ought to stick to the terminology already in use within psychology or develop radically new terminology wholly disconnected from the terminology now in use. Proposals along these lines have been advocated by both Machery and Millikan in their own ways.³⁰³

Machery proposes that psychologists and philosophers cease the use of “concept” entirely, replacing such talk with talk directly referencing prototypes, exemplars, and other representational kinds common to psychology. The term “concept,” he claims, fails to pick out a natural kind, as the representations that fall within its extension (prototypes, exemplars, theories, etc.) fail to have sufficient uniformity to support the sorts of systematic inductive

³⁰³ See Machery (2009), pg. 241-243, and Millikan (2017), pg. 46-49.

generalizations needed for the kind to be natural. As such, the term fails to pick out a kind that is likely to be useful for the explanatory purposes of psychology. My own view would seemingly run afoul of this argument as well insofar as I hold that it is still useful to use a general term, “conceptual structures,” to cover all of these coarse structural concept kinds. If “concept” fails to pick out a natural kind, my own “conceptual structures” would seemingly fail to the same.

Millikan meanwhile argues that what philosophers have termed “concepts” do not exist at all:

Putting this all very harshly, my claim is that there are no such things as empirical concepts.⁴ Classical empirical concepts are theoretical entities posited by a mistaken theory and must go the way of phlogiston and the bodily humors. What are left, I will be arguing, are merely names, things named, and individual people’s idiosyncratic unicepts for these things, which are supplied with information by equally idiosyncratic unitrackers.³⁰⁴

Concepts were supposed to be essentially public entities, singular things graspable by many distinct minds, and no such entities exist. She instead advocates that we talk of idiosyncratic unicepts (her atomistic concept alternative) and the unitrackers which guide their application and usage (roughly, her equivalent to Fodorian sustaining mechanisms, i.e. cognitive mechanisms that, in part, serve to setup and maintain the mind-world relation that determines the representation’s referent). Millikan also thinks traditional concepts erred in other ways. They were taken to be of classes, but unicepts must be able to represent particulars

³⁰⁴ Millikan (2017), pg. 49.

and other ontological categories in thought. Classical concepts were thought to be immune from error, redundancy, and equivocity, all of which Millikan's unicepts can suffer from (and by design). And concepts were thought to be discrete and clearly separable, whereas unicepts can be continuous with and overlapping one another. In other words, classical concepts exhibit a variety of features that Millikan thinks representations with their role should not have. Unicepts, lacking these features, are thus better fit to play that role.

While each approach has much to recommend it, I endorse neither. Machery's proposal has two marks against it. First, Machery's claim that "concept" fails to name a natural kind is poorly motivated. The various kinds of representation Machery endorses have much in common and it has long been productive to theorize about them under a common heading. They all factor in processes underlying core cognitive capacities and, if I am right, they are all constituted by the same kind of atomistic representation. I fail to see why that is enough for them to constitute a natural kind. While they may not possess the same inductive potential as prototypes and similar representations, they do factor in high-level explanations of cognition and there is a good deal that can be said about how they function. As Georges Rey in his Review of Machery's 2009 book argues, "Concepts seem to be natural kinds at least to the extent that they are the kinds of entity over which psychology generalizes."³⁰⁵

Second, on my view eliminating the use of my other general term "conceptual atoms" would leave us without any consideration whatsoever of the constituents of PA-supporting representations. Machery argues that a theory of concepts is not in the business of offering an explanation as to how we can bear such PAs to aspects of our environment, and thus is not in

³⁰⁵ Rey (2009).

the business of explaining the constituents of thought. But this is overly dismissive. Nearly all philosophers and many psychologists do see this as part of the theoretical goals in developing a theory of concepts, and without good motivation for abandoning such a goal we should attempt to accommodate it if at all possible.

On Millikan I have less to say, in part because her theory is far more revisionary and its connections to more traditional theories of cognition harder to chart. However, her main argument that concepts do not exist is based on the same type/token identity confusion that misled Frege into identifying senses with transcendent, abstract entities.³⁰⁶ Publicity in thought does not demand token identity between concepts; it does not demand that there be any one thing that both of our thoughts are related to. It demands, at most, that our two distinct, token thoughts are of the same type (and by extension that they are constituted by type-identical constituents). Both what I have termed “conceptual atoms” and “conceptual structures” can meet these demands and thus be public in the needed sense. The mere fact that concepts are public does not give reason to suppose that they do not exist if what it means to be public is properly understood.

Regarding Millikan’s other worries, this may come down to a simple disagreement over what we take to be essential to a more traditional understanding of concepts. My conceptual atoms are much like her unicepts and my conceptual structures are, in many ways, similar to

³⁰⁶ See Laurence & Margolis (2007) for a summary of the arguments by Frege and various Fregeans for the view that concepts are public, abstract entities, as well as reasons for thinking these arguments all hinge on confusions about token/type identity for concepts and overly simplified ways of thinking about concept individuation.

networks of unicepts on Millikan's view.³⁰⁷ My CAs can represent more than just kinds, representing particulars, properties, events, and whatever other ontological categories about which we can think. My CSs often overlap with one another, in part via their sharing of CAs. A single CA can be core or peripheral relative to what CS one is considering. The cognitive architecture of the mind on my view is a tangled web of CSs all intertwined with one another. And finally, as I have been at pains to stress throughout, there is no requirement that CSs be adequate or their possessors competent. And when a CS does prove adequate, that adequacy will always be relative the purposes of the possessor and the environment they are embedded in. It may be a merely linguistic dispute, but I suspect "concept" can survive these revisions.

These concerns about eliminativism are, however, largely pragmatic, having to do with what are productive and unproductive ways of using language in scientific and philosophical contexts. I am open to revision of our use of words like "concept" and terms related to it, such as "conceptual atom" and "conceptual structure," in such theoretical settings. Perhaps there are more productive terms for referred to what I call by those names. All that I insist upon is that whatever way we end up talking about conceptual representations, it be a way that is consistent with the existence of and the essential division between the constituents of thought and those representations which guide our engagement with the world and support our propositional attitudes regarding it.

³⁰⁷ This is not a coincidence, as the view I developed here originated as an attempt to connect Millikan's unicept/unitracker theory to more traditional theories of concepts from psychology. An argument could well be made that I would be better off adopting her terminology and abandoning talk of concepts entirely.

5.5 – Takeaways and Potential Lessons

I will now end this chapter by considering a few issues and implications that may follow from the view presented here. One parting concern may be that this all seems a roundabout and rather circuitous route just to arrive back at what appears to be suspiciously close to the simplest statement of atomism that appeared all the way back in chapter 4. After all, conceptual atoms seem to have just ended up being what one of the simplest statements of atomism held concepts to be, i.e. conceptual representations lacking in semantic structure and acting as the constituents of thought.

This, however, is not the case. It is true that the simpler statements of atomism posit concepts that are more or less identical to what I call here conceptual atoms. But what is distinctive of my view, I think, is the relation between these representations (whatever we may call them) and the various representations that have been posited by psychologists and given the name “concept.” Atomists, Fodor especially, always saw prototypes and the like as competitors, alternative proposals for what concepts may be like. It was considered a major failing that the representations posited by psychologists did not behave in the same ways as conceptual atoms, especially regarding the difficulties they faced with composition and thus accounting for cognitive productivity and systematicity. And psychologists for their part saw atomism in a similarly antagonistic way. Atomism was seen as implausible and explanatorily impotent, unable to explain anything concepts are supposed to explain, and tethered to an absurd and extreme form of concept nativism. It was seen as view to be dismissed offhand rather than seriously engaged with.

On my view, all of this debate was mistaken. It may be that CSs such as prototypes cannot compose in the ways necessary for productive and systematic thought. However, given

that they are not thought constituents this is hardly a knock against them. Likewise CAs, being fully without semantic structure (individuating or otherwise), are incapable of supporting core cognitive capacities and assisting their possessors in their practical activities. They are inert and, in isolation, entirely useless. But that is to be expected. They are simple building blocks out of which useful conceptual structures can be constructed. It is part of their function that they combine with other atoms in productive and exploitable ways. To expect CAs to do the work of CSs is akin to expecting an individual brick to provide the kind of shelter from the elements that a house affords.

The debate between philosophers and psychologists regarding the explanatory role for concepts has always depended upon the assumption that philosophers and psychologists were talking about the same thing. Without that assumption, the debate dissipates entirely.

Finally, I believe that the view I have presented here provides a few simple lessons and recommendations for those involved in the intractable, decades long interdisciplinary debates over the nature of concepts. For philosophers, it may be advisable to stop criticizing psychological theories on the grounds that what they call concepts could not serve as the building blocks for thought. It is far from obvious that psychologists are even in the business of providing a theory of such things, despite their occasional assurances to the contrary. Likewise, it may be best for psychologists to stop criticizing philosophical theories (especially atomism) on the grounds that their representations cannot do the work expected of the concepts

posited by psychologists. It is again not obvious that philosophers are in the business of providing theories of such things.³⁰⁸

It is my hope that the view presented here allows for the unification of both philosophical and psychological approaches to conceptual representation, suggests an effective division of labor between both fields, and may serve to harmonize their often competing assumptions and goals. If this is too great an ask, then I hope it at least promotes greater consideration of just how the various views of philosophers and psychologists relate and encourages more attempts at mapping those relations in novel ways.

I believe a second lesson is that talk of *the* concept of anything should be treated with great suspicion. If my view is correct, then in the vast majority of cases there is no such thing as “the” concept of anything.³⁰⁹ This is because, first, in many if not most cases there will be multiple coarsely structured CSs corresponding to a particular aspect of the world. Thinkers can have prototypes, exemplars, and theories that all represent the same thing in differing ways. There is also nothing stopping the development of new kinds of coarsely structured CSs, whether through intentional creation in ourselves or in the artificial psychologies we may someday create. Such new CS kinds may arise due to anything from the complex unguided

³⁰⁸ Why, for instance, would we expect that the specific representations by which a thinker accomplishes some task could be discerned without careful examination of the thinker in its natural environment, along with perhaps a careful examination of its various neural processes? Absent a few dubious assumptions about the relation between our intuition and our conceptual structures, there is little reason to think that reflection alone can provide us with knowledge of the fine structure of our own conceptual structures. And even granting those assumptions there is no reason at all to suppose it can provide such knowledge about the fine structure of the concepts of others, whether human or non-human animal.

³⁰⁹ This will also be the case under high-level atomism, so it does not depend on the truth of the specific view that I have advocated for in this chapter.

processes involved in biological evolution to the training by software engineers of a complicated algorithm underlying some artificial intelligence.

Second, there is a high degree of variability in terms of the fine structures of conceptual structures as well. Distinct prototypes of one and the same thing may vary in fine structure according to differences in experience, culture, upbringing, interests, ends, and so on. And a single prototype can have various fine structures across an individual's conceptual development, changing to reflect their varying experiences and needs over time.

Worse than the simple failure of such monolithic concepts as “the concept of ...” to exist is their potential for abuse. It has long struck me how appeals to “the concept” of something can act as kind of cognitive cudgel against less common ways of thinking. Appeals to “the concept” of something in philosophy (and likely other fields) often involve the elevation of one's own idiosyncratic and narrow conception to a universal status it does not deserve.

Even talk of “the expert concept of x” and “the concept of x among ordinary Americans,” which seems comparatively innocuous, is problematic insofar as there is rarely if ever any serious investigation into the matter.³¹⁰ In practice, what often seems to occur is that a philosopher reflects on the inferences they themselves are disposed to make, or perhaps considers in addition those of their peers, and holds that those inferences are the one's characteristic of the concept in question. But what reason is there for supposing one could generalize from the features of one's own concept to those present across a population of millions? Or even generalize to such a conclusion from a small population of mostly homogenous academic philosophers with heavy overlap in terms of assumptions and

³¹⁰ X-Phi, whatever its other merits, at least takes seriously the task of determining the structures of the CSs present in a given community and utilizes methods fit for that task.

theoretical commitments? It is cognitive anthropology from the confines of the armchair, and absent the assumption that concepts do not vary greatly within a given population (an assumption I have gone to great lengths to show is mistaken), there is little reason to suppose that any success can be had by these methods.

Appeals to “the concept” of something, or of the conceptual truths and necessities it supposedly underwrites, of the failure of those who disagree to even successfully think of what is under discussion, are I suspect often attempts (likely unintentional) to insulate one’s philosophical claims and arguments from empirical refutation by ferrying them to a conceptual realm where such evidence bears no relevance. It is a way of shifting from talk about the real world to talk of what are supposed universal (and presumably accurate) ways of thinking about that world, the latter accessible a priori and with comparatively little effort, the former far more difficult to access and requiring great care. Perhaps with cognition, and with the rest of the world in which we find ourselves, we should avail ourselves of methods that demand a deeper engagement than reflection alone can provide.

6 – The Need for an Externalist Semantics for Conceptual Atoms and Structures

This chapter marks somewhat of a transition point for my project in this dissertation. From this point onwards, I will be concerned less with issues of cognitive architecture and the explanatory roles had by the various conceptual representations that have been posited. Rather, here and in chapter 7 I will begin to lay the foundation for an integration of the framework I developed in chapter 5 with a teleosemantic approach to representation.

My concern in this chapter in particular is not to go into the details of teleosemantics or how that approach could be applied to my own view. That I will save for chapter 7. Rather, my goal here is to argue that many of the concerns covered in earlier chapters, such as the need for CAs to be public and thus sharable, demand an externalist semantics (whatever form that semantics may take). It is I believe only a semantics of this kind that can properly account for both the wide variability of CSs and the possibility of their constituent CAs nevertheless being shared across thinkers. Later in chapter 7 I will make a case that a teleosemantic approach in particular is the approach most likely to succeed in accommodating this desideratum and others related to it.

6.1 – From Concepts to Conceptual Atoms and Structures

As we near the end, it is worth taking stock of what has been covered up until now. In chapter 1, I argued that concepts as traditionally understood by most in cognitive science are mental representations that perform two functions fundamental to the life of any thinking thing. They (1) function as the constituents of the more complex conceptual representations which support the having of various propositional attitudes, and they (2) function to make possible and guide the exercise of core cognitive capacities such as identification/categorization and

reasoning of many sorts. A variety of theories have been developed in psychology of what such concepts and the processes they factor in are like, such as the classical, prototype, and theory-theory, with each of these being surveyed in chapter 2.

None of these theories, however, appears to be explanatorily adequate when considered in isolation from the others. Each type of concept encodes a specific kind of information (prototypes for typical or cue-valid properties, definitions for necessary and sufficient properties, theories for causal, functional, and essential properties, etc.), and the exercise of many thinkers' mental capacities draws upon information of each of these sorts. Thus, in chapter 3 I argued that what is needed is a pluralistic theory of concepts which recognizes that they come in a wide variety of structured forms. Concepts of varying structural kinds are drawn upon by thinkers in a way sensitive to the thinker's context, stage of cognitive development, perceptual capacities, practical ends, and the means available to them to achieve those ends. For any given aspect of the world, there will not be a single concept of it that will suffice for the needs, purposes, abilities, and contexts of use for all possible possessors of a concept of that world-aspect.

This conceptual pluralism is only possible provided a flexible account of concept individuation, one that allows for high-level referential concept types (ones individuated in terms of what they represent) to be realized by a wide variety of lower-level concept types individuated in terms of conceptual structure at some level of abstraction. Molecularist accounts, whether they be inferentialist or causal/conceptual role, are unable to provide such flexibility. By individuating concepts at least partly in terms of their structure/inferential role, they effectively identify these high-level referential types with specific lower-level structural

types. To put it in other terms, on molecularist views any concept has the particular structure that it has essentially. Concepts thus cannot maintain their identities across structural variations.

Due to this, such theories cannot allow for the wide variety of concepts that thinkers seem to draw upon without reducing the molecularist position to a triviality that even most atomists can accept. And in part because of this inflexibility, molecularist theories place overly stringent conditions on the possession of concepts. On such views, possession of a concept of some world aspect *W* requires not just the possession of a concept that represents *W*, but in addition possession of a concept that is structured in such a way that it disposes one to draw certain inferences or have certain beliefs regarding *W*, or possession of a concept that occupies a certain causal role within the thinker's mind.

This assumption that concepts are essentially structured in particular ways has the result of requiring for concept possession that one's concepts be adequate for some task or limited range of tasks.³¹¹ If possessing a concept of *W* requires possessing a concept with a particular structure, and if in addition this structure (as is often assumed) leads the possessor to have beliefs and be disposed to make judgments that render one competent at engaging with *Ws*, then possessing a concept of *W* requires one to be competent at engaging with *Ws* in the relevant ways.³¹²

³¹¹ Exceptions may be cases of empty concepts, e.g. PHLOGISTON, where there is no competency to be had, i.e. no true judgments or sound inferences to be made. Though even here there may still be some sense in which someone can be competent in its use. One possibility is that there are certain social interactions and related desires that depend on the empty concept being structured in useful ways, e.g. effectively conversing with others about the non-existent chemical substance.

³¹² It is usually further assumed that the beliefs and judgments that render one competent at engaging with something are beliefs and judgments that are true. This assumption could be denied. See Hoffman (1998), ch. 8, Hoffman (2003), and Prakash et al. (2000) for arguments

On such thinking, it may be the case that possession of a concept of horses, to take one example, requires that one be able to reliably recognize such animals under at least normal conditions (however these conditions may be understood). But both folk and scientific psychological explanations often require the attribution of concepts to those ignorant of or mistaken regarding the things they think about, sometimes even systematically mistaken. Explaining the mistakes of a young child who categorizes on the basis of superficial rather than deep properties, or the mistakes of an adult with very few or largely false beliefs about what they think of, requires they first possess the concepts they misapply and which support their mistaken beliefs and frustrated desires. Basic explanations of this sort, e.g. that a young child categorizes a painted horse as a zebra because they falsely believe that what makes something a zebra is merely that it looks like one, require that the child possess concepts of horses and zebras that dispose them to make these false judgments.³¹³ Explanations of this kind cannot be accounted for by molecularists without allowing for a plurality of minimally structured, often inaccurate concepts, a move that would render the core claim of molecularism compatible with atomism and collapse the distinction between both views.

Due to these two issues (widespread conceptual variability and the general ignorance of many concept possessors), I argued in chapter 4 that we ought to instead adopt an atomist understanding of concepts. It is not agreed upon, however, exactly how to understand the core claim of atomism: that concepts are unstructured. In chapter 4, I argued that atomism can be

that perception can be useful without being accurate. I will, however, be assuming throughout that competency requires something like truth or accuracy.

³¹³ In this particular case, the issue is likely not so much that the child has a concept with a dysfunctional structure so much as they are applying that concept in contexts or for purposes for which it is not suited. Likely what is happening is that the child is applying a prototype for skunks representing only its superficial properties, and using this in abnormal cases where the superficial properties fail to be reliable indicators of skunkiness.

productively understood as the view that a concept of a world-aspect *W* can be realized by a wide variety of structured concepts, each of which represents *W* in differing ways. I argued as well that the structure of a concept can be considered at different levels of abstraction, corresponding to how definite the manner is in which that concept represents its referent as being. Along these levels will be many explanatorily irrelevant structural concept kinds, concepts which represent what they represent in bespoke and idiosyncratic ways. Such concepts may even be ad hoc, possessed by only a single thinker for a short span of time. But in some cases regularities will arise among concept possessors, due to a variety of mechanisms like shared methods of acquisition, such as species-specific innate acquisition mechanisms. This then has the result that some of these structural types will constitute natural kinds of the sort that are relevant to general explanations in both folk and scientific psychology.

In chapter 5, I argued that while atomism so understood improves upon molecularism, it still falls prey to a mistaken assumption about concepts common to most theories: that the two functions concepts are held to perform (to function as building blocks of thought and to guide and support the exercise of core cognitive capacities) are both performed by a single kind of mental representation. The primary issue with this assumption is that the concepts posited by psychologists do not appear to have the same functional or theoretical role as concepts as understood by philosophers. For philosophers, concepts compose to form the more complex representations supporting our propositional attitudes, and those attitudes in turn partly explain the behavior of thinkers, including how they reason about and categorize the things around them. The concepts of psychologists have the functional role of these latter, more complex PA-supporting representations; they function to store information about the environment that can be exploited by the thinker to guide their behavior. They do not appear

to function as building blocks of further representations in the theories of psychologists, as one would expect from how philosophers describe concepts.³¹⁴

Because of this mismatch in functional role (among other, more secondary, issues), I argued that there is no one kind of representation that can do all that has been expected of concepts. In addition, I argued that the word “concept” is ambiguous, its meaning differing from a philosophical to a psychological context. What are called “concepts” by philosophers are the mental representations which in combination compose what are called “concepts” by psychologists. The concepts posited by philosophers I termed, following from atomists like Fodor and Millikan, “conceptual atoms” (CAs), referentially individuated and semantically unstructured conceptual representations that function as building blocks for thought, and the possession of which makes possible thinking about what the conceptual atom represents. And the concepts posited by psychologists I termed “conceptual structures” (CSs), complex conceptual representations that function to guide the exercise of core cognitive capacities. Thus, CAs are the unstructured representations that compose to form CSs, including prototypes, exemplars, theories, and the like.

This view has, I argued, various advantages over a more traditional picture of concepts, and even over a more revisionary picture such as that given by conceptual atomism as I understood it in chapter 4. Despite this, what I have presented is less a theory of conceptual representations than it is a framework or a foundation for one. This is due to the fact that it

³¹⁴ Psychologists do sometimes talk about concepts being building blocks of thought, and concept combination is a major focus for many. There is much that can be said about this. However, the manner in which psychologists think concepts combine does not appear to be much like how philosophers do, particularly with them not requiring that concepts be compositional in the philosopher’s sense. See Murphy (2002), ch. 12 for an overview of how psychologists think about these issues.

leaves a major gap that requires filling: exactly how CAs acquire their identities, and with it an account of the content (both referential and cognitive) of CSs. There is a major gap between the identity of a complex conceptual representation and what/how its constituents represent their referents.

There is, I think, very little that can be said about concepts in isolation from considerations of how their content is determined, including issues of concern throughout this dissertation, such as the conditions for concept possession and the scope of conceptual variability. There has been much theorizing about concepts while assuming their identities, these often arrived at via intuitions regarding concept attribution in philosophy, or via the assumption that the concept retrieved in an experimental task is of whatever the word prompt refers or via various behavioral criteria.³¹⁵ This is opposed to an approach focused first upon exploring the explanatory role for concepts (or CAs and CSs in my case) and determining the appropriate theory of content determination given that role.³¹⁶ In Chapter 7, I will attempt to take just such an approach.

³¹⁵ See Malt et al. (2011) and Hampton (2020) for discussion of how these assumptions underly the methodology of psychologists studying concepts.

³¹⁶ Worries about the use of intuitions in debates about intentionality are not new of course. See Millikan (2000) and Fodor (2004b) for discussion of how these worries may relate to Kripke and Putnam style thought experiments. See Machery (2017), especially ch. 1-3, for extensive discussion of how results in the concepts and x-phi literature may give reason to be suspicious of any philosophical methodology that gives intuitions a central place. I suspect intuitions of category membership may, however, have a role to play in a study of concepts in the same way that intuitions of grammaticality have a role in generativist linguistics. As intuitions of grammaticality may tell us about our syntactic representations, so intuitions concerning categorization may tell us something about the representations, processes, and mechanisms underlying our cognition. But as with grammatical competence, the representations, processes, and mechanisms may end up looking quite divorced from what we would pre-theoretically expect given the simplicity and immediacy with which those capacities are exercised.

6.2 – The Need for an Externalist Semantics for Conceptual Representations

There are two features of CAs and CSs (and of conceptual representations more generally, however they are understood) that any theory of content determination will have to account for. The first is that there appears to be widespread conceptual variability, both in terms of the coarse and fine structure of CSs. Second is that the explanations in both scientific and folk psychology require the possession of CAs even by thinkers who know very little about the world-aspects represented by those CAs or who even have mostly false beliefs about those aspects.

Thus, there seems to me to be two desiderata for any plausible theory of content determination for CAs and CSs:

1. The theory allows for and explains how a wide variety of CSs with very different coarse and fine structures can nevertheless represent the same aspects of the world.
2. The theory sets possession conditions for CAs that are meetable by typical thinkers, who often have few and/or largely false beliefs about the things their CAs represent.

As discussed in chapters 3 and 4, molecularist theories, whether internalist or externalist in nature, struggle to accommodate both (1) and (2). And for reasons similar to those that have been covered extensively for the case of language, internalist theories of conceptual content are unlikely to satisfy (2). If the inferential relations that a CA stands in determines what it represents, then those relations need to be robust enough, accurate enough, and of the right kind that they determine the appropriate extension for the CA and its corresponding CSs. Or if instead it is something akin to a mental description encoded by a

CS's structure that determines the referential content for associated CAs, then that structure must again be robust, accurate, and of the right kind. But this is something that none of the CSs posited (outside of definitions in idealized cases) will deliver.³¹⁷

Consider prototypes as one simple example of this. Suppose that prototypes determine their extensions in an internalist manner, i.e. the extension of a prototype is the set of whatever world-aspects have the properties the prototype's cognitive content represents. The cognitive content of a prototype of some world-aspect W will, by its very nature, exclude atypical Ws from its extension and include non-Ws that merely resemble typical Ws in often superficial respects. The judgments made by those thinkers (such as very young children) who seem to rely solely on prototypes and other similarity based CSs, reflects this. It is the reliance on such similarity-based CSs to guide categorization that leads children to erroneously conclude that, for instance, changes in the mere appearance of a world-aspect can bring about a change in the kind to which it belongs. Or which would, if adults relied on such CSs in a similar way, lead them to categorize an object identical to a nickel in all intrinsic respects except for being a foot in diameter as a nickel nonetheless.³¹⁸ Possessing a high number of heavily weighted typical or cue-valid properties of a kind does not, as a general rule, make something belong to that

³¹⁷ Definitions, while capable in principle of delivering highly determinate referential content, will only deliver such content provided that they are themselves robust and accurate. Given the general unavailability of definitions, even for trained experts, it is unlikely that the actual definitions had by most ordinary thinkers will meet this demand.

³¹⁸ This is adapted from Rips (1989), who found that subjects judged that a circular object three inches in diameter is more similar to a quarter than a pizza, but more likely to be a pizza than a quarter. What this seemed to show was that subjects rely on something more than mere similarity when categorizing (at least in some cases), likely drawing upon their knowledge of the standardized size of quarters and how this is a result of an institution such as the US Mint. Despite the greater similarity of the object categorized to a quarter, it was nonetheless not categorized as such because a nickel's size is (for the moment) essential to it. Thus, the prototype and exemplar theories could not account for all of how people categorize and thus could not be entirely adequate theories of concepts.

kind.³¹⁹ Thus it should not be a surprise that the cognitive content of a prototype will fail to determine the proper extension for that prototype and its relevant CA in an internalist manner.

It may be argued that these issues can be alleviated by making the structure of the CS more complex. Perhaps all CSs are hybrids, with any CS being an amalgamation of two or more of the many other kinds of CS posited by psychologists. While a prototype or exemplar alone will fail to deliver a suitable extension (as will any CS relying on mere similarity), perhaps a hybrid CS composed of a prototype and a definition, an exemplar set supplemented by a theory, or all of the above, will fare better. Perhaps the cores of such hybrids, usually taken to be definitions or theories, will prove better suited to deliver the needed extensions for both the hybrid CSs and their constituent CAs. In other words, perhaps a CS like a prototype of some *W* represents *Ws* because it is a constituent of a more complex, hybrid CS that includes a definition for *Ws*, with that definition determining the extension for the hybrid and its component CSs.³²⁰

³¹⁹ There are at least a few kinds that, arguably, do work this way, namely those where the members of the kind of bound together by a kind of family resemblance. The names of some kinds even make this basis explicit. For video games, it is common for certain genres to have names in the form of “*x*-like,” where the *x* is a highly influential game. For instance, roguelikes are a genre of game whose uniting feature is that they are all similar to the 80s PC game *Rogue* in some central respects, e.g. the use of procedural generation for content and especially harsh penalties for failure, and the genre of *Metroidvanias* consists of games sufficiently similar to *Metroid* and/or *Castlevania*.

³²⁰ The possibility of hybrid CSs does raise a problem for my view insofar as it blurs the lines between CSs and CAs somewhat, in that it brings the view closer to more traditional views of concepts. It may be that the definition is a CS itself, and thus the hybrid CS contains another CS as a constituent. CAs then start to look like to be nothing more than the primitive concepts of more traditional theories. This is somewhat worrisome. However, there still remains, I think, two relevant differences between my view and more traditional ones. First, the typical CSs of psychology, including hybrids, are identified with PA-supporting representations, rather than being parts of them or something ancillary to them, e.g. as sustaining mechanisms for the concepts that compose to form PA-supporting representations. Second, contra Fodor’s

This addition, however, does little to alleviate the problem. For one, while augmenting prototypes and similarity-based CSs with definitions would help deliver determinate contents for those CSs and their core CAs (provided those definitions are both accurate and sufficiently robust), there is little evidence that such definitions are available and even less that thinkers possess knowledge of them for the majority of the kinds they think about.³²¹ Nor is it clear that all kinds even have correct definitions available for thinkers to know in the first place. The Classical Theory of concepts, e.g. the view that concepts are definitions, has well-known and (as of yet) seemingly unresolvable problems.³²² To the extent that hybrid CSs incorporate definitions as cores, they inherit all of these problems.

Similarly, the principles and attributed essences found in folk theories tell their possessors more about the persistence conditions for kinds than what falls within their extensions. Knowing that a dog is a dog because of its doggy ancestry can tell you that a dog will not cease being a dog simply because it comes to no longer look like one. But knowledge that it is ancestry that matters (by itself), while useful for judging the sameness of a particular animal over time, does little to tell you, in context, which animals are of that kind in the first place. For that you would need to know in addition what any given animal's evolutionary ancestry is, and that is not something we are ordinarily in the position of knowing prior to knowing the identity of the animal itself.³²³ Thus, knowledge of theories (like with definitions)

compositionality requirement for concepts, it need not be the case (and I believe it is not) that the cognitive content of CSs be a function of the CAs that compose them. Rather, it may be that the referential identity of a CA is at least partially variable depending on which CS of which it is a part is being considered.

³²¹ Fodor et al. (1980).

³²² See ch. 2 for discussion of some of these issues.

³²³ Presumably the order is the reverse of this: I infer that an animal came from dog parents because I believe that it is itself a dog, and that only dogs can produce dogs via reproduction.

cannot help either because the kind of information encoded by theories is neither robust enough nor of the right kind to determine the proper extensions for their constituents.

Second, and more importantly, the shift to a hybrid theory does nothing to eliminate the problem of ignorance and error. To the extent that a thinker's definition or theory of a kind is mistaken or impoverished, the same issue will present itself in a renewed form for the hybrid. When individuals think they know a definition for a kind, they are usually mistaken (as the history of attempted analyses in philosophy should make clear). The definition of knowledge a justified true belief stood for over two-thousand years despite its now well-known flaws. And many believe the chromosomal definition of biological sex despite the correct definition used in the biological sciences defining sex in terms of gamete production. Likewise, the folk theories of individuals represent things as having placeholder essences, rather than determinate ones, or make use of false but accurate enough shortcuts, such as reasoning about causal relations between world-aspects via CSs that exploit constant conjunction.³²⁴ To the extent that the hybrid CSs or the CSs that comprise them include false information, they (and their core CAs) will fail to have the appropriate extensions under internalist theories.

A breeder may, perhaps, know an animal is a dog via its ancestry rather than its appearance (one could imagine a breeder who knows their dogs only as names on a ledger). But even then, at some point, whether with that dog's parents, its grandparents, or even further, the initial categorization of those animals as dogs was made on the basis of superficial properties. The positing of deep, essential features happens only after one is already thinking of a kind that one has mentally latched onto via recognition of its superficial properties. We similarly did not discover a kind with the chemical structure H₂O and then later infer that such stuff is a clear, odorless liquid filling lakes and rivers (though such an order is, in principle, possible).

³²⁴ Gopnik et al. (2001). Importantly this is not the only way that children, even very young, e.g. 3-4 year olds, infer causal relations between things. They also make use of more theoretical knowledge to infer the presence of such relations. But as with adults categorizing things on the basis of superficial, perceptually available properties, the existence of a seemingly less "sophisticated" method of reasoning does not preclude the possibility of more elaborate method, nor does it show that they are ignorant or somehow mistaken.

These issues are, I suspect, insurmountable for internalist theories, for many of the same reasons that similar issues in the case of language are. While there is more that can be said on this point, these issues are well-trodden ground. The simplest way to bypass the problem is to simply deny the linkage between the cognitive content/structure of a CS and what it and its CAs represent. CSs provide a particular perspective on what they represent, representing it as being a particular way. But so long as that perspective is able to be inaccurate or impoverished (or like a prototype is not even functioning as a perspective meant to capture the determinate boundaries of a kind), the contours of that perspective cannot be what determines what it is a perspective on.

6.3 – Externalist Semantics and Conceptual Variability

Given these issues with internalist theories, what is needed is an externalist theory for the content of CSs and CAs, some theory that will offload some of the unmeetable demands of internalist theories from the thinker onto their environment. However, the mere fact that a theory is externalist in nature does not entail that under that view (1) CSs admit of any degree of structural (whether fine or coarse) variability, or (2) CAs have possession conditions which are consistent with attributions of concepts to the uninformed or confused, whether for giving scientific or folk psychological explanations of others' cognitive capacities and behavior. In other words, not just any externalist theory can meet the two desiderata for a theory of conceptual representations laid out near the start of this chapter.

The reason for this is that under an externalist theory of mental content, the possession conditions for a CS or CA become largely a function of the conditions for a CA to have the referential content that it does. If on some externalist theory the condition for a given CA to represent some W is that the CA stand in some relation R to W, then possessing a CA of W

requires that one have a CA that stands in R to W.³²⁵ What is required to possess a CA of W then depends in large part on what relation R turns out to be. It also depends on the means by which a thinker can come to have one of their CAs stand in R to W and the degree to which those means are available to that thinker.³²⁶

Consider a very simple causal theory of mental content for CAs, one which holds that a CA of W represents Ws by virtue of the CAs of W being reliably caused by, and thus being causally correlated with, Ws. Such a relation is insufficient, as any representation will be reliably caused by a wide variety of things it plausibly fails to represent. This is why more sophisticated causal theories impose various restrictions on this relation. For instance, Fodor early on considered the requirement that the tokening of the representation happen only under normal conditions (however these may be understood), while later requiring that all non-content-determining causal correlations asymmetrically depend on the content-determining causal correlation.³²⁷ Another example can be seen with Dretske appealing to functions acquired by representations during normal development as a result of some kind of training or

³²⁵ This is, I take it, a largely trivial point. It is meant, however, to echo one of Fodor's main points about concept possession: that in providing an account of concept possession you thereby provide an account of concept individuation, and vice versa. If you know what it is for something to be a P, you know what it is to possess a P; and if you know what it is to possess a P, you know what it is to be a P. Likewise, if an externalist theory says that a concept of W must stand in relation R to W, then you need to possess a concept standing in relation R to W in order to have a concept of W (compare having a nickel and what one has being a disc of metal, produced by the US mint, intended for circulation, etc.). This issue, and its importance, is discussed in some detail in Peacocke (2009).

³²⁶ Largely, but not entirely, because part of what makes something a concept of W is that it is a concept in the first place, and what makes something a concept is plausibly entirely internal to the mind, e.g. its causal-functional role.

³²⁷ See Fodor (1984) for the attempt utilizing normal relations and teleology, and Fodor (1987, 1990a, 1990b, 1990c) for the attempt relying upon asymmetric dependence between causal laws. Note that Fodor's asymmetric dependence theory involves causal laws rather than correlations.

learning procedure, with their function being to be caused by and thus retrieved in the presence of what they represent.³²⁸ But while attempts to specify the right sort of relation is a major concern for causal theories and a source of much dispute, it will not matter for my purpose. I merely need an example of a simple content-determining relation for CAs to stand in for R, and to illustrate the point that not all such relations are fit for meeting the two desiderata I have set out.

Suppose in addition to the simple view laid out above that the causal correlation between the relevant CAs and Ws comes about via the CA's retrieval from memory (and with it the subsequent use of its connected CSs) being reliably caused by what the CA represents. And that this retrieval of the CA is guided by some subset of the CSs that it is a component of. It is because of this causal correlation that the CA represents W, and thus is a CA of W at all. And it is because it is a CA of W that the CSs of which CA is core will be CSs of W. A subject's prototype of wasps, to take one example, represents wasps and not bees as typically hairless and capable of stinging repeatedly because on this view the prototype's core CA represents wasps, and the core CA represents wasps because the prototype, among other CSs, leads to its reliable retrieval upon encounters with hairless, repeatedly stinging insects, thus setting up and maintaining a causal correlation between the relevant CA and wasps as a kind.

In order for a CA to be causally correlated with Ws its connected CSs must render their possessor able to, at minimum, reliably retrieve the CA from memory when confronted with Ws. Were a thinker regularly in the presence of swarms of wasps without ever retrieving a CS of wasps from memory (meaning they never identified nor reasoned about wasps despite their

³²⁸ Dretske (1981, 1986, 1988).

presence), or regularly retrieved such CSs when confronted by honeybees and houseflies, the relation between those CSs and wasps would fail to obtain. And without that relation being in place, what CAs and CSs the thinker has cannot, according to the simple causal semantics I am considering here, represent wasps. In other words, for a thinker's core CA and thus CS to represent wasps, that CS and others must at minimum be able to guide the identification of wasps well enough that the CA is usually retrieved only when wasps are present.

The result of all this is that something I have been attempting to avoid throughout this entire dissertation: the tying together of conceptual competency and concept possession. Up until now I have been arguing that this is primarily an issue that flows from molecularism and internalism. But here, the problem is simply reintroduced in externalist guise.³²⁹ If the content-determining relation between CA and W is something like mere causal correlation, then a thinker must still be disposed to draw the right inferences about W (those which establish the relevant causal relation) in order to have a CA of W. And this requires CSs to be both robust and accurate enough to reliably guide that CA's retrieval. This inferential requirement is not because such inferential relations are constitutive of the CA (as they are on molecularist/conceptual role approaches), but rather because it is only by being disposed to

³²⁹Note though that Peacocke is both a molecularist and an externalist. This is because he holds that some of the inferential relations which define a concept will make irreducible reference to world-aspects outside of the thinker, e.g. one of the inferences one must find primitively compelling in order to possess BLUE may be the inference from a blue thing being presented in perception to a judgment that said thing is blue (with the content of the perception itself externally determined in some manner).

and drawing those inferences that the CA bears the relation to W necessary for representing it.³³⁰

Despite the inability of many thinker's CSs to reliably guide CA retrieval in the way needed to establish a causal correlation, these CSs may nevertheless be entirely adequate for the purposes towards which they are put. This is especially so if the thinker is embedded in a sufficiently supportive environment or has relatively simple ends. The classic example of this problem is the predator-detecting mechanisms of rabbits, which are mostly inaccurate yet are highly adaptive and useful to rabbits despite this fact.³³¹ The rabbit, being quite low on the food chain and surrounded at all times by potential causes of its own death, keeps itself alive by overestimating the number of encounters it has with predators. The mechanisms, perhaps CSs, in the rabbit that function to identify predators produce more false positives than accurate identifications. This means these representations are unlikely to be causally correlated with actual encounters with predators, as opposed to events of bush rustlings and unknown sounds in the night. But for the rabbit it is better to have a hundred false positives than a single false

³³⁰ There may, however, be more than one set of inferences that can setup and sustain this relation between CA and what it represents. This will help with the problem of conceptual variability, as a CA's identity is a matter of its relation to the world, rather than the varied inferences or causal relations that may bring about that CA-world relation. It may not, however, help with the problem of ignorance and error, depending on whether the relation given by the externalist theory requires competency in one's engagements with the relevant aspects of the world.

³³¹ It may be objected that the mechanisms responsible for a rabbit's detection of predators is not conceptual, being perhaps a purely perceptual or even non-mental matter. While this may be true, it is not an issue here for two reasons. First, the specifics of the example matter less than its form; the main point is that representations that function guide the exercise of capacities like categorization need not be reliable in order to be useful to their possessors. And second, if the representation is used to guide the rabbit's recognition of predators, it becomes by definition conceptual as I am understanding it, and it is not uncommon for theories of concepts to limit or even collapse the cognitive/perceptual divide, such as in empiricist views of cognition (whether those views hold that concepts are perceptually formatted or that the conceptual primitives are perceptual in nature).

negative. The cost of the former is a small bit of wasted energy, while the cost of the latter is the rabbit's life. But even describing this mechanism and associated representation(s) functioning as a predator detector requires that they represent predators, despite the mechanism's unreliability/representation's general tendency towards inaccuracy.

Analogous cases in humans are not difficult to find. These have shown up throughout, but additional examples will help reinforce the point. Those with red-green colorblindness, specifically protanopia and deuteranopia, can still possess CAs like RED and GREEN, despite the fact they cannot reliably identify the presence of either. They may still, for instance, wonder whether the shirt they are currently wearing is red or green, something requiring their possession of the requisite CAs.³³² If someone confuses two people with one another, such as two similar looking students in a class or a pair of twins, they may consistently misapply their CAs of each, their CAs being reliably caused by one of the pair while representing the other. I personally possess the CA for wasabi and various CSs for recognizing it, mostly representing superficial properties like its taste, smell, and look. I possess these representations even though the overwhelming majority of times I have categorized a substance as such I have been wrong.³³³ The CA I associate with "wasabi" is correlated with ordinary horseradish mixed with

³³² They may possess other means or other CSs for identifying colors that do not rely on them being able to visually discriminate between those colors. Regardless, even if this were not the case, their inability to correctly discriminate between these colors need not imply that they have no CAs of them, and thus are not able to even think of them. And explanations of much of their color-related behavior requires that they are in fact able to think of them, e.g. my friend vacillating between two shirts because of worries about whether one is red, and would thus clash with his pants, or is green, and would thus complement them.

³³³ This is because actual wasabi is exceedingly rare in the United States (and not much more common in Japan), as well as being prohibitively expensive.

mustard and yellow dye, and not wasabi. Despite this it represents the latter, not the former, hence why the restaurants I frequent are deceiving me rather than giving me what I desire.³³⁴

The point here is not that causal semantics as a general approach to intentionality is mistaken or unworkable, nor that causal correlations and the technical sense of information related to them are of no relevance to representation.³³⁵ The point is simply that you have to enter into disputes over which externalist theory is correct with some sense already of what specifically you are trying to explain. And that the sorts of explanations given in both folk and scientific psychology involve the attribution of CAs to thinkers who are unlikely to have representations that meet the conditions imposed on them by some externalist theories.

What I will argue is that a consideration of how CSs fit into the lives of their possessors, how they are used in the contexts normal to them and in service of their particular needs, points towards a teleosemantic approach as the most plausible. It is that approach that is best equipped to explain the functions of CSs and CAs, to allow for widespread variability between CSs, and to allow for the possession of CAs by those thinkers who may strike us as intuitively rather limited in their understanding of those world-aspects that they think about and engage with. Finally, it is because a teleosemantic approach can deliver these that it is best equipped to ground the sorts of explanations that CAs and CSs are meant to provide.

³³⁴ Recently I have updated my CSs of wasabi upon realizing just how widespread the mislabeling is in American restaurants (and thus how unreliable using both visual properties and the words written in menus are for identifying it). A more reliable means is simply to rely on taste, as real wasabi is both milder and complex in flavor than its imitations.

³³⁵ It is intuitively plausible, for instance, that representations of perceptual properties acquire their referents via some process involving correlation, such as their being produced by mechanisms that have as their teleofunction the production of representations that are so correlated. See Neander (2017) for a thorough defense of this position.

7 – Towards a Teleosemantics for Conceptual Atoms and Structures

In chapter 6, I argued that little can be said about conceptual atoms (CAs) and conceptual structures (CSs), including crucially both (a) how widely they can vary in structure from one another, and (b) what their possession conditions are, without first providing a semantics for both. Further, I argued that an internalist semantics for such representations is unlikely to accommodate the desiderata arrived at in earlier chapters. These include that:

1. The theory allows for and explains how a wide variety of CSs with very different coarse and fine structures can nevertheless represent the same aspects of the world.
2. The theory sets possession conditions for CAs that are meetable by typical thinkers, who often have few and/or largely false beliefs about the things their CAs represent.³³⁶

Thus, what is needed to turn the mere framework for CAs and CSs given in chapter 5 into an actual theory is an externalist semantics, one that can specify under what conditions a CA is a CA of some world-aspect *W*. This requires specifying what relation *R* needs to obtain between a CA and *W* in order for that CA to represent *W*. However, not just any relation will suffice. Different relations will impose different constraints on thinkers, and whatever relation *R* turns out to be will need to be one compatible with (1) and (2) above.

In the chapter that follows I will argue that a broadly teleosemantic approach is able to accommodate (1) and (2). I will start by briefly discussing what it would mean to explain the exercise of core cognitive capacities, as such explanations form an inseparable part of how

³³⁶ Keep in mind throughout that this is equivalent on my view to saying that they possess CSs that represent the referent of their core CA and do so in ways that are not especially robust or even inaccurate. On my view, to have a CS is just to have whatever beliefs and other PAs that CS supports.

teleosemantics understands the determination of mental content. I will then lay out some of the core notions distinctive of teleosemantic approaches, in particular that of a teleofunction and an exploitable relation. Next, I will give a brief description of a teleosemantic explanation of the exercise of a core cognitive capacity supported by a CS. I will then end the chapter by discussing some of the outstanding problems the view still faces.

The goal will not be to deliver a complete semantics for CAs and CSs, and thus not to deliver a proper theory for both. And much of what follows will be necessarily schematic and rough in nature. Therefore, what I offer in this chapter does not fill in the gap between the framework for CAs and CSs of chapter 5 and the theory of CAs and CSs one may hope for. It likely does not even come anywhere close. Doing so would be a research program unto itself, and likely not something any one person's work, whether of a lifetime or a single dissertation, would be wise to promise. Rather, it is meant to be the start of a more thoroughly worked out semantics for both CAs and CSs, one drawing upon already established ideas and theories from the teleosemantic tradition in philosophy. If the division of concepts into representational atoms and the structures they compose is the foundation for a theory, what I offer here is something akin to a house's frame sat atop that foundation. It is lacking in detail and definition. But it does give a sense of the shape the final product may someday take.

7.1 – Externalist and Teleosemantic Explanations

As laid out all the way back in chapter 1, one of the main goals for a theory of concepts is to explain the various core cognitive capacities, both the having of them and their exercise. This goal is retained for my own view with CSs, as they are meant to explain the same thing. What would it mean to explain the having and exercise of a core cognitive capacity though?

One way to think about this is in terms of the judgments that a thinker in fact makes, and to take the task of psychology to be the explanation of why thinkers make these judgments as opposed to others they could have. This is the approach that many in psychology take:

In psychology, the study of concepts has focused almost entirely on the individual, internal level . . . With its interest in psychological representations, most of the psychology of concepts has simply said little about any way of specifying concepts apart from individuals' judgments.³³⁷

Some psychologists even define the extension of a CS as either what a subject believes to fall under that CS or what they judge to fall under it.³³⁸ On this approach, the explanandum is why the subject judges as they do. There is no consideration of the correctness of those judgments. And there is no consideration of whether these judgments facilitated the thinker satisfying their needs or achieving their ends. Nor does it seem that in these experiments, owing to the artificial conditions under which they take place, that the subjects studied aim at any end beyond successfully following the instructions of the experimenter.³³⁹ In addition, the things

³³⁷ Murphy (2020). Note that while psychologists may not recognize my term “conceptual structure” (though some do use similar language), I mean by it just whole class of representations that psychologists have called “concepts.”

³³⁸ See Hampton (2020). Hampton seems to use these two understandings of “extension” interchangeably, which may indicate that beliefs are being understood as or taken to perfectly align with the judgments of subjects. Note though that Hampton's two ways of defining extension are not equivalent. Someone can make a judgment contrary to what they believe, e.g. I may truly believe the death penalty is immoral in all cases, but in a passionate moment (say after watching a true crime documentary) I may judge that it is justified in that case. Conversely, I may believe that something is the case while failing to judge so in the moment (or even making a negative judgment) due to performance considerations, e.g. tiredness, inebriation, etc.

³³⁹ That is not to say that this way of thinking is universal, even among psychologists. The experiment by Posner and Keele (1968) mentioned below does consider the accuracy and inaccuracy of subject's judgments, to take one example.

being categorized are often themselves artificial in nature, being sorted into similarly artificial categories, such as the early experiments by Posner and Keele that focused on the categorization of dot patterns.³⁴⁰ Thus the study is of what judgments thinkers actually make regarding often artificial kinds, under artificial conditions, and in service of artificial ends.

This approach has two interrelated problems. The first of these is that it disconnects the study of the structure and use of CSs from the ends their possessors use them for. The desire to dutifully carry out the instructions of an experimenter does not approximate either the typical nor the normal desires that thinkers have in their day-to-day lives, nor can it adequately stand in for the wide diversity of form and content such desires exhibit. The second is that it leaves no room for subjects to be mistaken in their judgments. If the subject takes some world-aspect to fall under one of their CSs, it does (sometimes even as a matter of definition), or is at least treated as such for purposes of the experiment.

Why though should this matter? What is wrong with the explanandum that these psychologists have set for themselves? On this matter, I agree with Nicholas Shea, who argues that this explanandum can be accommodated with a purely causal, nonrepresentational theory. And if this is so, then we ought to dispense with psychology entirely (and prototypes, exemplars, theories, and all conceptual structures and atoms along with it). We should admit defeat and learn how to do neuroscience instead, or at least recognize that what we are producing is at best a temporary waystation on the path to an actual true theory of behavior.

Imagine our goal is to explain why the subject says “yes” when given the prompt “Is the object displayed on the screen an apple?” or extends their hand to hit the button marked

³⁴⁰ Posner, Keele (1968, 1970). The exemplar tradition in the concepts literature is particularly known for its use of highly artificial categories and equally idealized models of categorization.

“yes” when the image is flashed before them.³⁴¹ An explanation of this could, in principle, be given that talks of nothing over and above sound waves striking the ear drum of the subject and the complex intervening neurological events that eventuate in the subject’s vocal cords producing sound waves that realize a token of “yes.” Likewise an explanation could be given in such terms of how the pattern of irradiation of the subject’s eyes gives rise through complex intermediate neurological events to the raising and extending of the arm, following by the pressing of the finger into the relevant button. If our goal is to simply explain what does in fact happen in these cases, then any purported semantic properties of the internal states of the thinker are of little explanatory value, except perhaps as useful fictions in the short-term. To attribute such properties in an uncompromising manner would be to make the same mistake as attributing to anaerobic bacteria a desire to avoid oxygenated water when a purely physical explanation in terms of its magnetosomes and their relation to magnetic north/south would do. As Shea says of such explanations of mental systems:

A complete causal account of the operation of the system will be available in non-contentful terms. Proximal stimulation at input will cause the system to undergo various internal transitions that eventuate in movements at output. Intermediate entities at

³⁴¹ Note that the description of the explanandum here is not austere enough to fully capture what the psychologist is actually trying to explain if Hampton and Murphy are correct in their characterizations. The behavior of “saying ‘yes’” requires the movement of the vocal cords to be related to representational states, such as an intention to say “yes.” Behaviors, as Millikan (1993), ch. 7-8, notes, cannot be individuated in terms of the mere motions of the body and its parts through space. Were that the case, then a peacock performing a mating dance and one having a seizure would be engaged in the same behavior, provided that the two move in identical ways. Rather, behaviors must be individuated functionally, with those functions in many cases derived from the functions of the representational states that cause them.

various stages of this process may *have* semantic properties, but content does not figure in the underlying causal story about internal transitions and bodily movements.³⁴²

Instead, Shea argues that semantics shows its value in allowing for better explanations of the patterns of interaction between thinkers and their environments than are available on purely causal stories. And not just any patterns of interaction, but patterns of *purposeful* interactions, patterns of the kinds of interaction where the thinker is trying to accomplish something. It is, I think, by focusing on what a thinker is trying to do in context that we can start to see what an actual explanation of the exercise of a core cognitive capacity looks like.

Importantly, the two issues mentioned above (the disconnection of CSs from their possessor's ends and the elimination of the possibility of error) are also related insofar as a study of CSs that did focus on what thinkers were trying to do with their CSs would have no choice but to consider the question of whether the relevant CSs were being accurately applied and effectively used. This is because, in the normal case, correct applications of relevant CSs tend to lead to satisfied desires and achieved ends, while incorrect applications tend to lead to frustrated ones.

Consider a simple example. Suppose I am on a quiz show. As is normal, I have the desire to win. To win I must answer correctly, and I know this to be the case. Thus, I have a strong desire that my answers are the right ones. Suppose further that I have a CS of SONGS BY THE BAND KING CRIMSON, consisting of a set of exemplars for the songs of theirs that I know, and another CS SONGS BY THE BAND EMERSON, LAKE, AND PALMER, structured similarly as an exemplar set. Luckily (and quite unlikely), one of the categories in that night's game is 1970's

³⁴² Shea (2018), pg. 31-32.

progressive rock bands. However, when asked by the host what band the song Aquatarkus is by, I incorrectly answer that it was King Crimson. And I answer in this way because my exemplar for that song is a part of the exemplar set for King Crimson, rather than that for Emerson, Lake, and Palmer. It is not that, because I judge that Aquatarkus is a King Crimson song, I merely have a different CS of the songs of King Crimson with a different extension from those with greater prog rock knowledge. Rather, I am simply wrong, and in making my false judgment I frustrate both my desire to answer correctly and my further desire to win the game.

There is another reason to insist upon explaining purposeful environmental engagement rather than just actual observed behavior. It is I suspect, impossible to specify the extension of a CS without consideration of what a thinker is trying to do with it.³⁴³ This is hardly a new idea, especially among teleosemanticists. It runs throughout the work of Millikan and is embodied in things like Godfrey-Smith's Environmental Complexity Thesis, which states that the function of cognition is to allow a thinker to navigate the complex environments they are embedded in.³⁴⁴ I also suspect that it is the artificial contexts and the disconnect from a subject's desires that leads psychologists to understand things like a CS's extension in such a way as to rule out the possibility of error, or to define knowledge in such a way that it can be both false and unjustified, thus stripping it of its normativity. These themes will recur throughout the remainder of this chapter.

³⁴³ This is, I think, one of the core assumptions of teleosemantics, or at least those versions of it that focus on the use (or as is often said, the "consumption") of representations and how that use contributes to the performance of the functions of mental mechanisms and processes.

³⁴⁴ Godfrey-Smith (1996), pg. 3.

It is then, I contend, not enough for a theory of CSs to explain why a thinker T says “yes” or “no” when asked whether some world-aspect W falls under some kind K, or whatever other behavioral measures are utilized. The goal is not to explain what T *in fact* judges to fall under their CS, to include or exclude W and members of K from the extension of the CS depending on how T judges. Rather, it must be to explain how the CS allows for T to make the correct judgment that W is K under those conditions where so judging allows T to achieve one or more of their ends. In addition, it must explain what is going wrong when T judges that W is not K and how this frustrates their ends. A major goal in what follows will be to give an outline of what such explanations look like, drawing upon the resources and form of explanation characteristic of teleosemantic approaches.

7.2 – Teleofunctions for Conceptual Atoms and Structures

Before considering the specific sorts of explanations that teleosemanticists favor, it is first necessary to consider the notions that support and underly them. The core of teleosemantic approaches is the particular sort of function they posit, functions which differ in important ways from the more familiar functions in the philosophy of mind and psychology. The functions posited by teleosemanticists go by many names, including “proper functions,” “etiological functions,” and “teleofunctions.” For the purposes of this chapter, I will be using the term “teleofunction,” though I do not think there are substantial differences between the functions called by each of these names, at least as they are understood by many in the teleosemantics literature.

Before defining teleofunctions, one other notion will be useful: that of a reproductively established family (or as I will say, “reproductively established kind”).³⁴⁵ Such kinds are ones where each of the members are produced on the model of earlier generations of members through a copying process of some kind, with this being what grounds the uniformity and stability of the kind over time.³⁴⁶ What can count as a copying process is quite varied. Sexual and asexual reproduction are likely the most familiar cases, where offspring are partial copies of their parents or parent, with these reproductive linkages being what binds the species together. Reproductively established kinds exist for artifacts as well; Honda Civics form such a kind, as each year’s model is designed by Honda’s engineers partly via partly copying the design of the previous year’s model, e.g. the 97 Civic varies little from the 96 Civic, which varies little from the 95 Civic, and so on. In some cases, the copying is not from previous generations but from a common plan or template. A factory mass produces goods that are highly uniform, in large part because they are all copies of a common design.³⁴⁷

When considering a reproductively established kind, one can ask how that kind has persisted over time. Such kinds require an ongoing copying process for their continued existence. Humans would die out were members of the species to stop having children. The Honda Civic would eventually cease to exist if Honda discontinued the model. And any

³⁴⁵ See Millikan (1984), ch.1, for more detail on this and many other issues related to teleofunctions, such as how to define reproduction in the general sense, distinction between different kinds of reproductively established families, etc.

³⁴⁶ This is also what makes them constitute natural kinds in the sense outlined earlier in § 5.2.2. Members of a species share a great many properties because of the fact they are all partial copies of earlier generations of the species, and this sharing of many properties is what grounds the inductive potential for the kind as a whole and thus makes the kind natural.

³⁴⁷ Importantly something can belong to more than one reproductively established kind. Any given 1997 Infiniti Q45 is both part of the kind that includes the 97 and 98 Q45s and the kind that includes all other 97 Q45s, with the copying processes and properties shared differing between each of the two kinds.

consumer good can go the way of Pogs or Beanie Babies if the corporations producing them stops sending orders to the factories. And often these breakdowns in the copying processes are due to various selection pressures. Human beings may die out if an alien race discovers us and scours us from the face of the planet. The Honda Civic could grow stale in the eyes of consumers as competing models from Toyota and Ford win over consumers. And Pogs and Beanie Babies did, in fact, die out as consumers spent their money on other toys instead.

A member M of a reproductively established kind K has as its teleofunction the effect F if the current and continued existence of K and thus M is explained by ancestors of M producing F in the past. Put more simply, a thing's teleofunction is whatever effect(s) its ancestors produced that explains why that thing and the kind to which it belongs has survived.³⁴⁸

Teleofunctions are, I think, best understood by way of examples. The classic example of a teleofunction is the function of the heart. Hearts produce many effects that are incidental to their being hearts. They give off certain odors, produce bumping sounds, can serve as makeshift paperweights, and so on. That they pump blood seems more central and important to being a heart than these other effects, and a theory of teleofunctions can explain why. Hearts have kept their possessors alive via the pumping of blood, and by doing so contributed to the survival of their possessors as a kind. And it is because their possessors survived and

³⁴⁸ This definition is simplified in several ways. The full definition of teleofunctions, presented most in depth in Millikan (1984), ch. 1-2, is both highly complex and highly technical, and none of what will be discussed here will depend on those complexities. I suspect, however, that more detailed and comprehensive accounts of cognition will require the use of more of the resources that theories of teleofunctions have to offer, e.g. a theory of CS acquisition mechanisms would likely need to appeal to derived teleofunctions in particular, rather than direct teleofunctions (what I am calling just "teleofunctions" throughout).

outcompeted others that the genes which allow for the construction of hearts were passed on to future generations. Pumping blood, and not producing bumping sounds or other such effects, is what provided survival value to the organism, and in turn to the hearts housed within them. The pumping of blood is thus the heart's teleofunction because it is that effect which accounts for the survival and current existence of hearts as a kind.

Similarly, Honda Civics continue to be produced because consumers continue to buy them. And because consumers continue to buy them (and presumably because Honda executives like money), the Honda corporation continues to produce them. Consumers continue to buy them because Honda Civics do certain things for them, primarily transport them and their friends/family in relative comfort and economy from point to point. Were they to regularly fail to perform this function, e.g. Honda had several years where their cars were immensely unreliable, the kind may go out of production and go extinct. Thus, the teleofunction of a Honda Civic is to serve as transportation. Likewise, Pogs and Beanie Babies as a kind ceased to exist the moment they were no longer able to perform their teleofunctions of generating schoolyard clout and serving as dubious speculative investments respectively.

Another way of understanding what exactly teleofunctions are involves contrasting them with the main competing approach to functions, what I will follow others in terming "causal functions."³⁴⁹ Roughly, a causal function of some P is the effect produced by P that

³⁴⁹ Cummins (1975, 1983). See Millikan (1989) for commentary. Another kind of function that will not be considered here are so called "design functions," where the function of a kind is a matter of the intentions of its creator(s). Given that my concern here is primarily with naturally-occurring cognitive representations and systems, such functions are a non-starter for many CSs. They may have a role to play in cases where a CS is purposely created or altered, such as in purported cases of conceptual engineering. However, even in the case of artifacts I suspect teleofunctions provide a better account than design functions, though I will not argue so here.

causally contributes to the overall functioning of the system of which P is a part. In other words, it is the contribution made by P to the effects produced by that system as a whole.³⁵⁰ On this competing approach, a heart has the function of pumping blood, not because its ancestor's pumped blood and that the kind was perpetuated as a result, but rather because hearts in the here and now, in actuality, play the causal role of pumping blood within the organism of which they are a part. It is by pumping blood that the heart contributes to the circulatory system fulfilling its function to circulate blood and thus oxygen throughout the organism, and through this contributes to the functioning of the organism as a whole.³⁵¹

Causal functions suffer from at least one major problem not shared by teleofunctions: they lack an account of malfunctioning and may not allow for it at all. The causal function of a part of some system is what that part *in fact* causally contributes to the overall effect produced by that system as a whole.³⁵² If the system stops producing the effects that it typically produces, this does not mean that the system and its parts are malfunctioning, but rather that the systems and its parts have shifted in their functions. A heart clogged with plaque, one that can no longer pump blood, retains the teleofunction of pumping blood but loses the causal function of doing the same, as it is no longer producing this effect and thereby not contributing to the functioning

³⁵⁰This is the sort of function appealed to by classical formulations of functionalism, and by many appeals to causal-functional roles more generally.

³⁵¹Note, however, that theories of causal and teleofunctions do not need to be in competition. One could adopt a pluralist theory of functions (see Godfrey-Smith, 1996), such as holding that teleofunctions explain why a given thing exists, while causal functions serve to explain how a system and its parts function. Tellingly, Cummins does not see the inability to malfunction as a bug in his theory and is explicit about it having this feature: "...if something functions as a pump in a system . . . then it must be capable of pumping..." (Cummins 1980).

³⁵²It is the focus on the actual functioning of a part of a larger system that leads Aristotle to the odd claim that a hand ceases to be a hand when separated from the body. Once separate, the hand can no longer contribute to the functioning of the body as a whole, losing its function and thus its identity in the process. Aristotle seems to have something like causal functions in mind here, rather than teleofunctions. See Aristotle (2014), *On the Parts of Animals*.

of a greater system. And if a heart is defined by its function, then it is the heart's teleofunction and not its causal function that accurately characterizes it. Otherwise whenever one goes into cardiac arrest, the issue would not be that their heart is seizing up, but rather that they have ceased to have a heart, which intuitively is a far graver problem.³⁵³

By contrast, a thing fails to perform its teleofunctions when it fails to produce the effects that have historically accounted for the survival of its kind.³⁵⁴ Importantly, however, a thing can fail to perform its teleofunction in more than one way. In the case of an *internal* failure, the problem lies with the makeup of the thing itself. Lungs filled with fluid, such as in cases of pneumonia, are failing to perform their teleofunctions in this sense; the buildup of fluid makes it impossible to take air into the lungs and extract oxygen from it. On the other hand, were a person in the depths of the ocean where there is no air and no oxygen accessible to them, the problem is not with their lungs but rather with the environment external to them.³⁵⁵

³⁵³ In many cases, a thing's teleofunctions and causal functions align. Thus, in healthy individuals under normal conditions, the heart has both the causal and teleofunction of pumping blood. This is because both the effect produced by its ancestors and the effect it produces which contributes to the current overall functioning of the circulatory system (and thus organism) of which it is a part are the same effect.

³⁵⁴ This is admittedly too simple. Some things may never perform their functions without this being a failure. For instance, the gag reflex has the function of expelling blockages from the airways, but if lucky enough it may go entirely unused across one's entire life. When normal conditions are introduced, this can be made more precise, with failure defined as the failure to produce the effect when under normal conditions for its performance. For instance, the normal conditions for the gag reflex are when a blockage is in fact present, and the failure of the reflex to expel this blockage (and thus the probable death of the choking individual) constitutes a failure to perform its teleofunction.

³⁵⁵ Note that the environment external to the lungs need not mean the environment external to the organism in which they are housed. If the lungs cannot take in air because of a failure of the diaphragm to contract, such as in individuals with advanced ALS, the failure of the lungs is still an external one.

It is not a failure inherent to the lungs that they fail to pull double duty as gills. A failure of this latter kind is not an internal, but rather an *external* failure.³⁵⁶

The functions I require for CAs and CSs must provide some sort of standard for the normative evaluation of CA and CS functioning, and thus must allow for the possibility of malfunctions. As argued in § 7.1, the referential and cognitive content for a CS cannot be understood absent a consideration of their possessor's ends, as it is how a CS factors in an explanation of the satisfaction of those ends that gives us a principled means of assigning content to it. A CS can either perform this task well or poorly, and in some cases the poor performance of the CS will be due to external or internal failures of the CS.³⁵⁷ It is teleofunctions and not causal functions that seem best positioned to provide for the possibility of such failures, and thus it is functions of the former kind that I will draw upon in what follows.

A theory of teleofunctions brings with it two other useful notions that will be of aid in what follows. First, it tells you what counts as a normal explanation for the performance of a thing's teleofunction. A normal explanation is an explanation of how members of a kind have historically performed their teleofunction on those occasions where such performance contributed to the persistence of the kind in the face of selection pressures of some sort. The normal explanation for how the lungs take in air appeals to the contraction of the diaphragm to pull air into the lungs, rather than appealing to the operation of a mechanical ventilator that does the same work. Both accomplish the same goal, and the lungs successfully perform their teleofunction either way. But as our ancestors possessed diaphragms but lacked ventilators, it

³⁵⁶ As we will see, in many cases purported instances of inadequate or impoverished CSs are actually cases of external failure, cases where the CS is being used for ends or within environments abnormal to it.

³⁵⁷ In some cases it will simply be poor luck that block the functioning of a CS.

can only be the former that accounts for the historical persistence of lungs as a kind, and thus only the former can factor in an explanation of how the lungs operate in the normal case.

Second, a theory of teleofunctions tells you what counts as normal conditions for something's performance of its teleofunctions. This is a major advantage, as specifying these conditions is, in any case, a difficult proposition. Most things can only perform their teleofunctions within an incredibly narrow range of conditions.³⁵⁸ Human lungs, for instance, work only within a minority of environments on a single planet in a vast, intensely unaccommodating universe. For the teleosemanticists, the normal conditions for the performance of a thing's teleofunctions are those conditions that are referenced in a normal explanation for such performance. Equivalently, normal conditions are those conditions which obtained during the performance of that function, by ancestors of that kind, on those occasions when such performance contributed to the persistence of the kind in the face of selection pressures of some sort. The normal conditions for human lungs are those conditions where the air is a mixture of roughly 78 percent nitrogen and 21 percent oxygen, a barometric pressure of roughly 14.7 pounds per square inch, and so on. Slight changes to either, such as the decrease in barometric pressure near the summit of Everest, dramatically reduce the normal functioning of the lungs, if not outright prevent it.

Each of these notions (teleofunctions, normal explanations, and normal conditions) are ultimately needed to provide a teleosemantics for CAs and CSs. What I propose then is that we approach the question of what CAs and CSs represent by first looking at their teleofunctions.

³⁵⁸ This is, in part, why normal conditions cannot simply be understood as mere typical conditions. As Millikan has pointed out, typical conditions in our universe are conditions of a total vacuum with a temperature of near absolute zero, and even on earth conditions are more often hostile to our existence than amenable to it.

From there we can transition into looking at what explanations of CA and CS functioning involve, and how such explanations can provide an answer to what is involved in CAs and CSs not only succeeding, but oftentimes failing, to do what they are supposed to do. And lastly, we can utilize the notion of normal conditions to explain how the normal performance of a teleofunction can so heavily depend on a CS's possessor being embedded in a supportive environment, i.e. the normal conditions for that CS's functioning, and thus how a CS that looks inadequate in the abstract can be highly adequate in context.

It is important to make a few clarifications before moving further. The first is that there are more selection pressures than just natural selection that can be relevant to something acquiring a teleofunction. Artifacts, to take one example, can acquire teleofunctions as a result of the purchasing and usage decisions of consumers. A claw hammer has the teleofunction of driving nails because it is by driving nails that hammers have continued to be bought and used by consumers. It is because claw hammers continue to be bought and used that they continue to be produced by manufacturers. And it is because they continue to be produced that the kind continues to persist in the face of the pressures posed by often fickle consumers, e.g. other goods vying for their attention and money. Similarly, words, syntactic structures, and (most relevant to my purposes here) CSs can persist because their possessors find them useful, and perpetuate them in various ways, both individually and communally, e.g. continuing to use them and thus helping to keep them in memory, the explicit teaching of their usage to others (especially in educational contexts), reinforcement via culture and media, and so on.

Second is that a thing can have multiple teleofunctions. A claw hammer does not merely drive nails but extracts them from the material into which they have been driven. Claw hammers could have failed to have this second function, being like staplers which can staple,

but which standardly require a separate tool to separate staple from what is stapled.³⁵⁹ Likewise, a prototype functions not just to guide categorization but to guide reasoning and support linguistic comprehension as well. Whenever a kind persists because, historically, members of that kind produced a variety of effects, the kind will have multiple teleofunctions corresponding to each of these effects.³⁶⁰

Finally, a thing's teleofunctions can vary according to context and change over time. They may even be lost entirely under the right conditions. Regarding contextual teleofunctions, a piece of camouflage, whether natural or artificial, has the broad teleofunction of helping one blend into the environment. But it also has a narrower teleofunction of helping one blend into the specific environment that one finds oneself in at a given moment, e.g. blending into a forest, a desert, a cityscape, etc.³⁶¹

A change in teleofunctions can be found in the toy Play-Doh, which was originally sold and used as a wallpaper cleaner for the first twenty years of its existence, and it was this usage that accounted for its continued manufacture and existence at that time. Were Play-Doh entirely ineffective at this task, consumers would quickly grow tired of it and search for a new

³⁵⁹ Arguably, this could be a case of claw hammers simply not existing as a kind, and some other sort of hammer and a separate tool (say a nail puller) existing instead and occupying the same functional role in construction and repair that the claw hammer currently does.

³⁶⁰ Even in the case of kinds that seem to have a single, central teleofunction, there are actually a wide variety of teleofunctions possessed, though not all of these will be relevant to the sort of explanations one is trying to provide. For instance, while the heart's teleofunction of pumping blood is one of its teleofunctions (and the one it is often defined by), it also has various downstream effects that also account for its persistence over evolutionary time. The heart also circulates oxygen throughout the body by pumping blood, and by circulating oxygen it does (among many other things) keeping oxygen flow to the brain and keeping it alive. See Millikan (1984), pg. 34-38, for discussion of this issue and ways of determining which teleofunction is of most importance in explaining a given case.

³⁶¹ Functions of these latter, narrower sort are termed by Millikan "derived proper functions." See Millikan (1984), ch. 2.

product, leaving its manufacturer without any reason to continue producing it. However, it lost this function at some point during the transition from cleaning tool to children's toy, acquiring a new teleofunction in the process (perhaps something like providing entertainment to young children, or since parents are those making the purchases, keeping children out of their parents' hair).

Finally, a loss of teleofunction can be seen in the many examples of vestigial organs in humans and other animals that once had teleofunctions, but now persist despite producing no useful effects for their possessors. The auricular muscles present near the ears of some humans, for instance, once had the function of orienting the ear to sources of sound in the environment. Now they are only able to, at best, allow their possessors to wiggle their ears, an effect that (outside of contrived cases) is of little evolutionary value. Similarly blind mole-rats retain their eyes despite those eyes now being covered over by a layer of skin, thus being functionally useless to the mole. Their eyes once served a purpose, having teleofunctions involved in helping the mole-rat navigate its environment, among others. But now, likely given a change from an above to a below ground environment at some point in the specie's history, their now useless eyes exist only due to evolutionary inertia.³⁶²

³⁶² Due to the fact that a kind's proper function can change over time, Millikan (1984), pg. 32, draws a distinction between historically proximate and historically remote teleofunctions ("proper functions" in Millikan's terminology). The historically remote functions are those effects produced by members of the kind that are further in the past that account for the current existence of the kind, while historically proximate are those nearer in time. The mole rat's eyes may have no teleofunction that is especially proximate, but they do have a historically remote function of providing the rat with information about its environment. Kinds in general will have a series of teleofunctions, from the most historically remote to the most proximate.

7.3 – A Teleosemantic Approach for Conceptual Atoms and Structures

With the notions of teleofunctions, normal explanations, and normal conditions, we can begin to sketch an account of how CAs and CSs acquire their referential content, and from that how CSs acquire their cognitive content.

As mentioned in the previous section, determining the extension of a CS requires understanding how that CS is applied in service of its possessor's ends. Due to this, the starting point for this account will be the teleofunctions for directive representations, e.g. representations supporting PAs such as desires.³⁶³ What is the effect produced by such directive representations such that they have kept around? I will be following Millikan in taking the teleofunction of such representations to be the causing (in part) of their own fulfillment.³⁶⁴ It is when directive representations, working in concert with descriptive representations such as those supporting beliefs, are satisfied that they do something useful for their possessors, something likely to keep their possessors (and them) around for a bit longer. For the sake of simplicity, from this point onward I will mostly eschew use of “directive representations” for talk of desires and use of “descriptive representations” for talk of either beliefs or talk of CSs like prototypes, exemplars, and theories. Thus, talk of “the teleofunction of belief” or “the teleofunction of desire” should be read as shorthand for “the teleofunction of the representations which support the having of beliefs/desires.” And “satisfaction of a desire”

³⁶³ As mentioned earlier in this dissertation, I leave it open that directive representations may form a subclass of CSs. Regardless of whether this is the case, I will assume here that, whatever they turn out to be, they are constructed of the same CAs that CSs are. It is at least possible, however, that directive representations are wholly different in their makeup than descriptive representations like the CSs I have focused on throughout, with coordination between them being possible because of harmony between both their content and their functional relations, rather than a sharing a parts.

³⁶⁴ Millikan (1984), ch. 3.

should be read as “the performance of the teleofunction of a directive representation underlying a desire.” In cases where greater specificity matters, I will talk of directive and descriptive representations instead.³⁶⁵

A toad’s desire that impels it to eat flies and other suitable prey does nothing to keep a frog from death unless that desire is at least sometimes satisfied. A toad that never satisfies its desire for food does not live long, and it takes its desires with it when it goes. Likewise, my desire to win at a board game only persists so long as its satisfaction remains possible. If I fall too far behind, I may jettison that desire and take on new ones, such the desire to just have fun talking with friends as we play, save myself from embarrassment, or even the desire to ruin the game out of spite by flipping the table.³⁶⁶

Directive representations, however, can only fulfill their teleofunctions (in a normal manner, i.e. in a manner that accords with a normal explanation) by working in concert with descriptive representations, these including the CSs we have focused on throughout. In more personal terms, desires can only be satisfied if one has beliefs adequate to guide one’s behavior towards bringing about the state of affairs the desire represents. Thus it is plausible that CSs have as a distal teleofunction to serve as a partial cause for the bringing about of directive

³⁶⁵ See Shea (2018), pg. 177-194, for more discussion of the difference between directive and descriptive representations and how to specify each in greater detail and formality than I do here. Note that Shea is mostly concerned with subpersonal representations rather than PA-supporting representations.

³⁶⁶ There is a distinction that is being ignored here for simplicity’s sake, namely that between the teleofunctions of the directive representations supporting desires in general and the teleofunction for a directive representation that supports a desire with some determinate content. Plausibly, directive representations of the latter kind have teleofunctions that are derived from the teleofunctions of the mechanisms that produce them, whatever these turn out to be. I will largely ignore this distinction in what follows, but see Millikan (1984), ch. 2, for more on derived teleofunctions and how they differ from the more direct teleofunctions I mostly focus on here.

representation satisfaction. More proximally, CSs fulfill this distal teleofunction via the performance of the functions focused on since all the way back in chapter 1, namely categorizing, reasoning about, comprehending linguistic utterances about, and so on, the world-aspects that one's directive representations concern. My exemplar of a close friend helps me to pick out my friend in a crowd, and thus contributes to the fulfillment of my desire to reunite with them. My prototypes of sugar, flour, eggs, and the like contribute to my desire to make and eat cookies by helping me to recognize those ingredients and avoid using ones that will spoil my dessert.

7.3.1 – Exploitable Relations

This of course leaves open the question of how CSs normally manage to perform these proximate teleofunctions of categorizing and reasoning, and thereby perform their more distal teleofunctions of helping to bring about directive representation fulfillment. The answer, in short, is that the processes which CSs factor in exploit the fact that the CSs component CAs (1) stand in exploitable relations to aspects of the thinker's environment, and (2) stand in relations to one another, these relations themselves exploitable.

The idea of an exploitable relation comes from the work of Peter Godfrey-Smith (2006).³⁶⁷ For my purposes, I will understand an exploitable relation (very roughly) as a relation between two or more world-aspects that must be appealed to in giving a normal explanation for the performance of some thing's teleofunction(s). The two kinds of exploitable relation I will consider here are (1) a causal correlation between two world-aspects, and (2) a structural correspondence between either (a) two sets of world-aspects, or (b) the internal

³⁶⁷ See as well Shea (2018), ch. 4 and 5.

relations between elements in two worlds aspects and/or the sets containing them. These notions as stated are fairly obscure, so I will now try to clarify them by way of a few examples.

7.3.2 – Exploitable Causal Correlations

Consider first a simple case of an exploitable relation involving causal correlation. A smoke detector has the teleofunction of warning its owners of the presence of fire. It is because they do this reasonably well that they continue to be bought and produced. To warn of fire, they obviously attempt to detect the presence of smoke that is causally correlated with fire, thus exploiting the correlation between the two. But they also exploit causal correlations in order to detect smoke in the first place, with the correlation exploited depending on whether the detector uses a photoelectric or ionization method. In the case of the former type, the smoke detector exploits the correlation between the presence of smoke and a difference in the intensity of light detected in the smoke detector's optical chamber, this difference caused by the smoke's scattered of that light. In the case of the latter type, the smoke detector exploits the correlation between the presence of smoke and a difference in current between two ionization chambers in the detector, with this difference caused by the smoke's ionization of the air within one chamber (open to the air) relative to the current in a reference chamber that is closed to the air. Finally, the detector's owner will (hopefully) exploit the causal correlation between the loud beeping of the detector and fire to get to safety.

Importantly, if considered in an unrestricted manner these correlations may fail to hold. Someone who regularly microwaves popcorn and has an ionization style smoke detector in their kitchen may have noticed that such detectors can give off more false positives than

accurate warnings.³⁶⁸ Does this not then show that ionization-based smoke detectors do not actually exploit such correlations? How can they exploit correlations that may not even exist?

The issue here is one similar to the problem that faces causal-informational theories of mental content (at least, ones that lack a teleological component). As argued in chapter 6, such theories struggle with cases where detection mechanisms (or representational mechanisms more generally) are adequate, despite being inaccurate enough that they fail to establish causal correlations between the states of the mechanism and what is detected/represented.

This is precisely the problem faced by our smoke detector. However, when attempting to explain the functioning of a smoke detector, we restrict ourselves to explaining how the detector functions *on those occasions that it succeeds*. In other words, we are trying to give a normal explanation of how the detector performs its teleofunction, rather than giving a causal explanation of how it performs its causal functions at any given time.³⁶⁹ The smoke detector's teleofunction is to detect and warn of fires, and a normal explanation of how it performs this function appeals to the detector's exploitation of the correlation between differences in the currents between two ionization chambers and the presence of smoke. What matters is not that the correlation holds under all possible conditions, but rather only that the correlation holds within conditions normal to the functioning of the detector. If the correlation fails to hold

³⁶⁸ This is one reason why this style of detector is currently being phased out, despite having advantages in a small range of cases (such as flaming vs. smoldering fires). They also, perhaps worryingly to some, contain americium-241, and are thus slightly radioactive.

³⁶⁹ Compare explaining (in the way Hampton, Murphy, and many psychologists do) the actual behavior of a thinker and the behavior of the thinker that aims at and achieves some end of theirs. We care about cases where the smoke detector succeeds when explaining its functioning, not any and all cases where it happens to go off. And we ought to care, not about whatever the thinker may happen to do in any situation, artificial or otherwise, but rather how their CSs succeed in doing what they are supposed to do on those occasions they do.

outside of these latter conditions, it makes little difference to the normal explanation being given.

7.3.3 – Exploitable Structural Correspondence

The notion of an exploitable structure correspondence between inter-CA relations and relations between world-aspects is more complex than that of exploiting causal correlations. As such, I will dedicate a good deal of time to explicating it here.

Take an example from Millikan.³⁷⁰ A film camera functions by focusing the light reflected off of its subject onto light-reactive film. If the camera is functioning normally, what is produced is a photo whose elements are isomorphic to the elements in the state of affairs that caused it. If the camera is malfunctioning, either due to internal failures (a cracked or dirty lens) or external failures (a lens cap left on), it will fail to produce a photo whose elements stand in relations which map onto the relations that the elements of its subject stand in.

Suppose, for instance, that I take a photo of a stoplight head-on. The stoplight consists of a few simple elements in equally simple relations to one another. There is a rectangular housing in which three round, colored lights are embedded. At top is a red light, yellow in the middle, and green at bottom. If my camera is functional and I have my settings correct, the photo I take should be such that the relations between the elements in the stoplight bear a structural correspondence to the relations between the elements in my photo. The red, yellow, and green circular patches appearing in the photo should be related in the same manner as the lights they represent are (red above yellow, yellow above green). The distances between the patches should be proportional to the distance between the lights, as should be the distances

³⁷⁰ Millikan (1993), introduction.

between the corners of the rectangle encompassing the patches and the corners of the stoplight's housing. It is because the relations between the photo's elements map onto the relations between the parts of the stoplight that I can learn certain things about the latter solely from the former, e.g. the sizes of the lights relative to their housing. In such a case I exploit the fact that the photograph preserves the spatial relations of the subject in miniature to come to learn about the spatial properties of that subject.

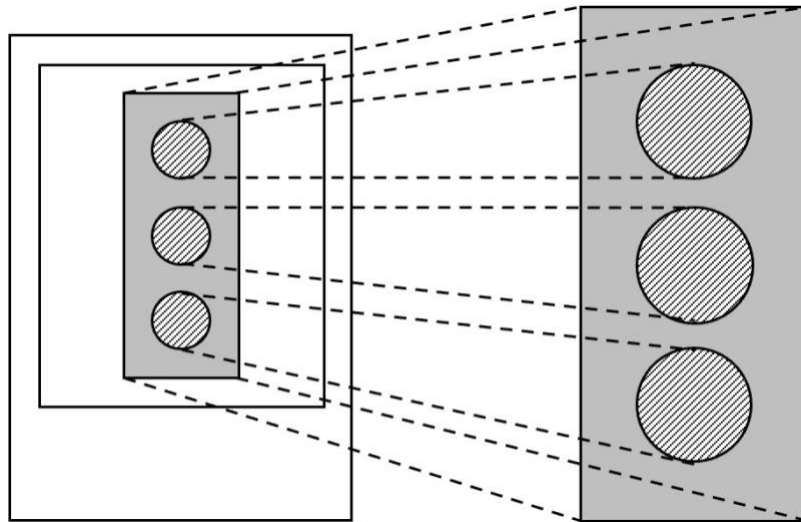


Figure 7.1

Importantly, the exploitable relations need not be purely spatial, as they are in the above case, nor do they need to be as complete and consistent as they are in the case of a photo. The lines connecting stations on a subway map are not proportional to the actual distances of the subway lines they represent. Nor do the paths the lines take on the map bear much of any resemblance to the paths actually taken by the subways. But the connections formed between stations and lines on the subway map will, if the map is a good one, map onto the actual

connections between stations and lines in the real world, e.g. the map will not have a line connecting stations A and B if no such line actually exists or omit a line connecting B and C if one does. And it is because of this that the subway map, despite abstracting away a good amount of detail, can nevertheless be useful to a traveler making their way through a city, exploiting the structural correspondence between the lines on the map and real world lines.

It may be argued that this places attention on the wrong part of the case, focusing on the manner in which the map is used rather than how it is produced and for what purpose. Imagine a case where a subway map has been found out of context and without its usual labels (perhaps due to age the text has come to be illegible). And suppose further that the map bears a structural correspondence to not just the nearby subway lines but, by pure accident, to a collection of hiking trails and rest areas just outside of town as well. The finder of the map notices this latter correspondence and takes the map to be of these trails rather than the subway lines. And if they go on to use the map on their hiking expedition, an explanation could be given of their successful navigation in terms of their exploitation of the correspondence between the elements of the map and the elements of the system of trails they are travelling.³⁷¹

There are two things to say about such cases. First, this seems to be a case where the teleofunctions of the map user's directive representations are being performed in an abnormal way. It is by pure accident that the representation they are relying upon (the map) happens to bear a structural correspondence to the world-aspect they are traveling (the system of trails).

³⁷¹ Cases like this are symptomatic of a much broader problem with relying on structural correspondence in theories of representation: any given world-aspect will have parts whose relations bear a structural correspondence to some set of world-aspects or their parts. In other words, there are too many structural correspondences, nearly all of which seem irrelevant to representation. Thus, there must be some principled means of selecting one or some subset of these relations that plausibly are explanatorily relevant.

In the normal case, maps (whether paper maps external to the mind or cognitive maps internal to it) are produced by mechanisms which have as their teleofunctions the production of maps with such structural correspondences. It is part of a normal explanation for the production of a subway map that it is made by the metro system to serve as a representation to be exploited by subway riders.³⁷² And it is part of the normal explanation for whatever cognitive mechanisms produce mental maps that they produce representations which bear structural correspondences with aspects of the world in some way causally connected with that mechanism.

Second, while the use of the map in this way may be accidental now, over time it could come to be part of normal explanations concerning engagements with the hiking trails, rather than with the subway lines. Suppose the hiker has great success in navigating the trails using this map, despite the correspondence between the two being entirely accidental. Suppose further that they then make copies of this map, perhaps adding a few labels here and there, and then distribute those copies to their hiking friends. Over time, the fact that the map is useful to the hikers comes to be what explains the continued production and persistence of the map, rather than the uses subway riders put it towards previously. The map then comes to acquire a new teleofunction. The use of it by the hikers (and their exploitation of its structural correspondence to the trails) becomes part of a normal explanation for how those hikers satisfy their desires to stay safe and have fun while hiking. The map, like Play-Doh before it, comes to be used for new ends, acquires new teleofunctions, and now factors in different normal explanations than before. A representation may accidentally help in the performance of a

³⁷² The relevant teleofunctions to be explained in the case of such an explanation are likely that of the intentions and other directive representations of the map's designers or those who work for the metro line and local government.

teleofunction on an occasion or two. But continued use for new purposes will shift what it is that the representation represents and the normal explanations it factors in.

7.3.4 – Normal Explanations for the Exercise of Core Cognitive Capacities

With the preceding notions laid out, we can now sketch a normal explanation for the exercise of a core cognitive capacity. In this section, I give the general form for such an explanation and then give an example of what explanations of that form may look like. For the sake of simplicity, the example will concern the exercise of a categorization capacity and will make use of a relatively simple prototype representation.

As discussed at the start of the chapter, CSs must be understood partly in terms of how they assist in the performance of the teleofunctions of directive representations. This is because a distal teleofunction of CSs is to act as a partial cause of the successful performance of such directive representation's teleofunctions. CSs do this via the performance of more proximate teleofunctions, such as supporting capacities like categorization and reasoning. Further, they are able to perform these proximate teleofunctions by encoding information that can be exploited by the cognitive processes underlying those capacities.

CSs encode this information in the relations between their component CAs. Thus, it is the teleofunction of CAs to stand in such content-supporting relations to one another and to aspects of the environment. These relations I termed, following Godfrey-Smith and Shea, exploitable relations. The two most important of these relations are (1) a CA standing in a causal correlation with some world-aspect, and (2) a set of CAs bearing a structural correspondence to a set of world-aspects. The cognitive processes underlying core cognitive capacities exploit these relations to guide a thinker's behavior towards the satisfaction of their ends, i.e. the normal performance of their directive representation's teleofunctions. And

perhaps most importantly, what those CAs represent is a matter of what exploitable relations they must stand in under the normal conditions referred to in a normal explanation for the functioning of the CSs of which they are a part, e.g. what world-aspects they carry information about and what set of world-aspects/elements a set of CAs stand in a structural correspondence relation with.³⁷³

A normal explanation for a CS supporting the exercise of a core cognitive capacity then appeals to how its component CAs stand in relations to each other and the world and how the processes underlying that capacity exploit those relations. The referential content of those component CAs is given by the exploitable relations they bear to the world under the normal conditions appealed to in such an explanation. And in turn, the referential content for a CS is inherited by its core CA, and its cognitive content inherited by the CAs forming its periphery.

With the general form of such explanations now laid out, we can move to consider a more concrete (if still schematic) explanation. Consider a case where someone is driving a car. They presumably (or at least hopefully) have desires to stay safe and to obey the rules of the road. This includes making sure that they stop at any stop signs they encounter along the way. Thus, they have directive representations which support these desires, and which have as their teleofunctions the serving as partial causes of bringing the car to a stop at each of these signs.

In order for these directive representations to do this, they must be assisted by CSs. The relevant distal teleofunction of these CSs in this case is to help in bringing the car to a stop at stop signs along the way, and the relevant proximate teleofunction is to ensure each such stop

³⁷³ Compare to the case of the subway map. The map has as its teleofunction helping subway riders to navigate the metro system. The normal conditions that must be appealed to in explaining how the map helps riders do this involves the lines composing the map standing in exploitable relations to the actual subway lines they represent.

sign is recognized in time. A wide variety of CSs differing in both coarse and fine structure could do each of these jobs.³⁷⁴ Suppose though that in this case the driver draws upon a prototype CS for stop signs, one with the following component CAs and weights:

Prototype: STOP SIGN		
CA ₁	Red	0.7
CA ₂	Octagonal	0.5
CA ₃	“Stop” written on surface	0.4
CA ₄	7 Feet Tall	0.3
CA ₅	White Rimmed	0.1

Figure 7.2

A person possessing such a prototype believes redness and being octagonal to be the most typical properties of stop signs, with this reflected in the weighting of their respective CAs. Because of this, it is these properties they rely on most when trying to identify the presence of such signs. In most cases, this likely works quite well owing to the structural correspondence between the CAs present in the prototype and the actual properties had by stop signs. The CAs found in the prototype stand in statistical relations to one another that correspond to the statistical relations between the relevant properties in the real world, at least provided the prototype is accurate in how it represents stop signs to typically be.

But causal correlations are likely being exploited here as well. The CAs composing the prototype have to be retrieved on the basis of something. Given that the CAs represent

³⁷⁴ Someone may, for instance, recognize it via the prior recognition of the word “stop,” while another (say someone who does not know English) relies instead on its shape, color, and height. And as we will see, stop signs are not uniform across time and place, and thus CSs with differing fine structures will each fare differently depending on where they are made use of.

primarily visually-available properties like redness and being octagonal, it is plausible that their retrieval is being guided by perceptual representations of these properties. These retrievals need to be reasonably reliable. The miscategorization of a color can, in many cases, result in severe consequences, e.g. a driver who mistakes green for red on a stoplight. One way to ensure that these retrievals are reliable is to exploit causal correlations between shape and color properties and the elements of the perceptual representations that represent them. Then the conceptual system(s) can exploit the plausible causal correlations between these perceptual elements and the CAs representing those properties.³⁷⁵

When the driver pulls up to the sign, the light reflecting off of it and into their eyes leads to the production of a perceptual representation that represents, roughly, an object of octagonal shape, red in color with a band of white along its outer edge, sat atop a long cuboid object anchored to the ground, at a certain distance and bearing certain relations to other objects in the surrounding area.³⁷⁶ These elements of the perceptual representation are in normal cases of perception correlated with the properties they represent. And it is this fact that the processes responsible for the retrieval of CA₁, CA₂, CA₃, and CA₅ exploits.³⁷⁷ These four CAs form part of the periphery for the prototype CS of stop signs, and the structural correspondence between

³⁷⁵ I assume here (I think plausibly) that perceptual representations of an object having some perceptually-available property tend to give rise to, and thus are correlated with, judgments that the object have that property under normal conditions for the relevant mechanisms.

³⁷⁶ I assume here that the perceptual representation is not representing high-level properties like being a stop sign or even having the word “stop” written on it (that presumably involves linguistic representations of some sort). It is not, however, essential to my argument that the representational scope of perception be limited in this way (though I am sympathetic to the view).

³⁷⁷ CA₄ may have its retrieval guided by the processes involved in linguistic comprehension, which themselves involve CSs. This is a complication I will ignore for present purposes, however. If needed, imagine that the driver is not a speaker of English and simply sees the letters as yet more shapes and colors represented in perception, with relevant CAs retrieved on that basis.

CA₁, CA₂, CA₃, and CA₅, and the properties of the stop sign are exploited by the processes that the CS factors in. The weights of these four CAs are summed, and the category threshold is easily passed. This triggers a positive categorization judgment of the object as a stop sign, which along with the directive representation to stop the car at such signs causes the driver to press on the brake and bring the car to a halt.³⁷⁸

In order for this normal explanation of the desire to stop the car at stop signs to proceed in this way, it must be the case that under normal conditions the CA₁, CA₂, CA₃, and CA₅ that compose the CS guiding the action that will ultimately satisfy that desire both (1) stand in causal correlations to the corresponding perceptual elements that guide their retrieval, and (2) bear a structural correspondence to the relevant properties in stop signs. It is in virtue of this that these CAs and the CS they compose each represent what they represent.

The prototype CS may, of course, fail to perform its teleofunction in the way laid out above, for both internal and external reasons. It may be, for instance, that the driver does not realize that some stop signs on private property and in Hawaii are blue, rather than red. Or that stop signs in Japan, contrary to nearly everywhere else in the world, are downward-pointing rounded triangles rather than octagons. Such failures are external ones, cases where their prototype CS is being used in an environment for which it is not adapted. Conversely, the driver could be tired or inebriated, in which case the retrieval of the CS's component CAs may be too slow for the recognition of the stop sign to be made in time. Failures such as these would

³⁷⁸ There is obviously much more than this going on. The driver, for one, must know that pressing the brake pedal slows down the car. And this knowledge is, presumably, supported by their folk theory of the car's operation.

be internal in nature; the CS is well adapted to the environment in which it is used, but its performance is in this instance diminished.

What matters though is what happens when things go right. What happens when the driver does successfully satisfy that desire to stop the car, and what must the CS and its component CAs that supports that satisfaction represent in the normal case to help bring that about. As mentioned multiple times before, this explanation is incomplete in many ways. It does not even begin to approximate the full complexity of such representations and processes. But I hope it does at least give a sense of their flavor.

7.4 – Advantages to the Approach

Throughout this dissertation I have argued that one way in which traditional theories of concepts err is in treating concepts in an acontextual matter. This treatment manifests in at least two ways. First is that there is often not enough attention paid to how accessible the properties represented by the concept are to the thinker, with there often being a focus on deep, defining or essential properties of things that are not easily identified by the thinker in realistic contexts. Second, is that there is a lack of appreciation for how concepts that may be inadequate in general can be highly effective given that the thinker has suitable ends and is embedded in an accommodating environment. In conjunction, these tendencies lead to a picture of concepts according to which they are highly complex and require a good deal of cognitive sophistication to possess and wield.

This acontextual understanding of concepts is, I think, implicit in monist approaches to concepts and talk of “the” concept of a thing more generally. If there is a single concept for a thing, then it would need to be the case that a single concept could be fit for all the tasks a thinker may put it towards and all the environments in which they find themselves. And if

concepts are what psychologists and many philosophers understand them to be, i.e. mental representations that (among other things) function to guide the exercise of capacities like categorization and reasoning of various kinds, then these tasks would need to be drawing upon the same information of the same sort, despite differences in the nature of the tasks or the contexts in which they are carried out.

My own approach attempts to avoid this. Thinkers often draw upon knowledge of bespoke and highly contingent properties when categorizing and reasoning about the things with which they interact. And many of these properties are not plausibly represented by a universal CS held by all those capable of thinking about what that CS represents. Additionally, oftentimes the CSs had by thinkers are, by the standards of the acontextualist, highly impoverished, representing very little if anything diagnostic of the aspect of the world being represented. And in some cases, CSs of the same coarse and fine structure may represent distinct things as having the same, highly undiagnostic properties depending on the needs and context of their possessors.³⁷⁹

These issues are clearest in cases of animal cognition. Returning to the case one last time, a toad will attempt to eat small, fast-moving objects and attempt to flee from large, immobile objects that it takes to be nearby. And these representational capacities are not especially flexible nor sophisticated. The toad will attempt to eat metal BBs shot near it and flee from sufficiently large pieces of square cardboard placed in its vicinity. It seems that

³⁷⁹ Putnam's Twin Earth case is a case of this sort, though less severe than some of the cases I consider. In that thought experiment, the average thinker uses superficial properties like clear, liquid, and the like to guide their usage of "water" (likely using something like a prototype) and the reliance on these same properties on Twin-Earth results in a difference in the referent of "water" (and plausibly in what is represented by the corresponding prototype).

whatever properties the toad is relying on in making its identifications are simple ones, such as those involving size, rate and direction of movement, and so on. Given the inflexibility of their recognitional abilities and how easily the toad is deceived, some have argued that the toads are only representing things such as small-black-moving-object and large-immobile-object in these cases, rather than things such as prey and predator.

On a teleosemantic approach to CS content, the inflexibility of the toad's recognitional capacities does not, in itself, show that the toad is unable to represent more abstract aspects of its environment. The toad's limited recognitional capacities reflect the uniformity of its normal environment and the relative simplicity of its ends. In the toad's normal environment, trying to eat small, fast-moving airborne objects is a highly reliable way to acquire food. Further, eating these flies is the toad's only fly-involving end; unlike humans, the toad has no interest in conducting entomological studies of flies or pinning them for purposes of insect collecting, and thus no need to represent anything relevant to either (such as representing the fly's thorax for purposes of proper pinning). Similarly, fleeing from large nearby objects is a good way for the toad to avoid getting eaten, and the toad has little interest in engaging with predators for any other reason than this, such as hanging out together by the pond. The breakdown of these recognitional capacities when the toad is placed in a university lab with professors shooting BBs and placing cardboard nearby is no more a consideration against the toad possessing CAs for prey and predator than the breakdown of human capacities for recognizing water on Twin-Earth is a consideration against humans having a CA for water.

Another case can be found in gosling imprinting. The young gosling needs to stay nearby its mother during the early stages of its life, both for purposes of food and for protection from predators. In order to do this, though, it must first identify its mother (or caregiver, food

provider, protector, or however else one believes the content should be cashed out).³⁸⁰ And as with the toad, the mechanisms and processes in the gosling seem to rely upon simple, undiagnostic properties to make this identification. All that appears to be needed is that the thing identified as mother be both a certain rough size and nearby to the gosling during a critical period of its development. Thus, it is possible for young goslings to imprint upon and end up closely following everything from the boot of a famous zoologist to a cardboard box mounted on a model train. And even a casual perusal of YouTube, Instagram, or TikTok will turn up countless examples of goslings, ducklings, and other bird young imprinting on barn cats, pet dogs, birds of another species, and countless other things.

As with the toad, the gosling's mechanisms for imprinting are easy to lead astray. However, that the gosling succeeds in imprinting only when its mother is in fact the large nearby object during that early critical period does not entail any inadequacy on the part of this mechanism. Whatever is responsible for goslings imprinting on their mothers works good enough when operating under normal conditions, conditions which include the mother and not mischievous zoologists with model trains being nearby during their early development. That it fails in highly abnormal conditions the mechanism was never designed to operate in does not show that the mechanism is not really for identifying and imprinting on mothers, rather than simply large nearby objects.

Human, gosling, and toad cognition are, I suspect, not as different from one another as may be suspected, at least not in their fundamentals. Human CSs are also often impoverished

³⁸⁰ It is I suspect (and will briefly argue) indeterminate which of these is represented. However, this is not a sort of indeterminacy that I suspect matters for explanatory purposes. The explanation of how the gosling survives its youth goes through whether it involves recognition of mother versus recognition of food-provider.

by the standards of philosophers, yet entirely adequate given the needs and contexts of use of their possessors. Given a supportive environment, CSs do not need to be especially robust to meet the needs of their possessors. My CS WATER would be wholly unreliable on Twin-Earth, with me likely constantly identifying XYZ as water. I could go my whole life on Twin-Earth bathing in, imbibing, and doing laundry with XYZ all while taking it to be water.³⁸¹ But even on Earth, under mostly normal conditions, my CSs for water are fairly limited, requiring a specific and highly contingent context for their reliable use. As a child I failed to distinguish between white wine and sparkling water, solely because I had not yet realized that water is not served in wine glasses. Were wine commonly served in drinking glasses, my competency at avoiding such mix-ups even as an adult would drop precipitously. The same is the case for a great many of my CSs; my ability to reliably identify my favorite cereal brand at the supermarket is always one mascot change away from being diminished. And my ability to reliably distinguish Coke and Pepsi is entirely dependent on me seeing the labeled containers from which they are poured, as I cannot distinguish them by look or even by taste.

Consider a more detailed example that may already be familiar: someone categorizing coral snakes in North America. Being able to identify a coral snake on sight is crucial to hikers as they are highly venomous and commonly encountered on popular trails. Luckily coral snakes have bright coloring that makes them stand out from other snakes and their environment, with alternating bands of red, yellow, and black across the length of its body. Less lucky is that many coral snakes look superficially similar to a variety of non-venomous snake species,

³⁸¹ This of course assumes that what my WATER represents does not shift during that period. While there is much debate on this issue (see Boghossian 1989 and Burge 1988), the question of whether the content of a representation would switch or remain the same is I think, much like with questions of possession, something that can only be answered within the context of a well-worked out theory for content determination for the relevant representations.

differing only in the ordering of their color bands. Anyone who hikes or has been in an organization such as the Boy Scouts may know some variation of the mnemonic for distinguishing the venomous coral snake from its similar, non-venomous doppelgangers, such as “red touching yellow, you’re a dead fellow; red touching black, you’re ok Jack.”

Categorizing snakes on a hike by looking for the order of color bands is a reasonable and mostly effective strategy: the ordering of the bands is highly correlated with the species of snake and the ordering can be discerned at a safe distance. Certain properties more highly correlated with venomous coral snakes would be more reliable in making such judgments, such as the shape of the snake’s snout. But given that these properties are less safely accessible, requiring one to get within striking distance of the snake to identify, they tend to go unused by all but herpetologists.³⁸²

While looking for color band order is a broadly effective strategy, it is one with limitations. In the Americas some coral snakes have atypical color patterns due to genetic abnormalities, and the further south one goes the less likely it is that the mnemonic fits the coral snake species one may encounter. And for old world coral snake species the mnemonic is entirely unreliable. So, a CS that represents color band patterns in coral snakes can be a highly effective means of recognizing and avoiding them, but that CS’s effectiveness is limited to a relatively small segment of the earth, and even in those contexts it is far from a guarantee that it will not lead you astray. But the usefulness and adequacy of the CS is in no way diminished if its user fails to recognize these limitations, and would without hesitation falsely categorize coral snakes in abnormal (for them) environments as safe. All that matters is that it

³⁸² Herpetologists are those who study snakes and reptiles.

works often enough and for enough people that the CS is useful to its possessors and it thus perpetuates itself. In the same way, it only matters to the toad and gosling that their CSs facilitate fly eating and mother following just often enough for those CSs to be selected for and persist.³⁸³

The takeaway is that CSs can be highly simplistic in their structure (and thus in how they represent what they represent) and yet still be entirely adequate for the needs of their possessors. It should not be denied that some thinker possesses a CA of some world-aspect solely because they are unreliable at recognizing or engaging with those aspects. It must first be considered where the relevant CSs are being used and towards what ends. What appears inadequate and unreliable in an abnormal context may be entirely adequate in the thinker's normal environment or adequate despite its relative unreliability. Teleosemantic approaches, as I have been arguing, can accommodate these facts. That, I think, is the strongest argument in their favor.

7.5 – Some Outstanding Issues

While the approach I advocate here is, I think, the best fit for the framework of CSs and CAs I have developed, it is not without its problems. Teleosemantic approaches in general face a range of issues, ones that my own view acquires by association. While I cannot here provide satisfactory responses to each of these, I can at least gesture in the direction of a few potential and promising solutions.

Perhaps the main objection that has been raised against teleosemantic approaches to representation is that they fail to deliver fully determinate representational contents. Under

³⁸³ I am assuming, as seems likely, that whatever mechanisms or representations present in the toad and gosling responsible for these behaviors are innate.

such approaches, what a representation or mechanism represents is given by what counts as normal conditions for its functioning, as given in a normal explanation for how its teleofunctions are performed. These normal conditions involve such representations (or elements of them) standing in exploitable relations to world-aspects in the thinker's normal environment and to one another. If this is so, then if there are multiple adequate normal explanations for the performance of their teleofunctions, there will be multiple possible world-aspects that are (or could be) represented by that representation and its elements. For instance, one normal explanation for the functioning of the mechanism involved in the toad's catching and eating of flies appeals to how that mechanism exploits the correlation between small, airborne, moving objects and flies within its normal environment. And it is through this that the mechanism can serve as a reliable means of identifying flies and guiding the behavior that results in their being eaten. But a normal explanation that appeals to food (or toad-food) works just as well as an explanation that appeals to flies.

CSs function to guide the exercise of capacities like categorization, and reasoning. It is by doing this that they are useful to their possessors, and persist due to evolutionary, personal, and cultural processes. A theory of CSs should explain how they manage to perform these functions, as well as the reasons why they sometimes fail to do so, and in so doing either facilitate or stymie the performance of the teleofunctions of related directive representations. In other words, a theory of CSs should explain when and how CSs are adequate for the needs of their possessors. Explanations of this sort do not seem to demand that representational contents be fully determinate, but rather only that any indeterminacy be sufficiently limited. The toad's representations being indeterminate between fly and toad-food does not matter; explanations appealing to either are near equal in terms of their quality. What would matter is

if those representations also represented things such as the states of the toad's brain/body, or stimulus patterns on the toad's retinas, that cause those representations. This is because explanations appealing to these aspects would be unilluminating.³⁸⁴

There will also remain CAs for which there is no obvious externalist account of content determination, teleosemantic or otherwise. Empty concepts remain an issue for any theory of concepts, and while teleosemantics can appeal to what a representation *would* represent under normal conditions that will never obtain, this answer remains incomplete. There is also the difficulty, similarly shared with any theory of concepts, that as what is represented grows more abstract it is less clear how a CA can come to represent it. One can tell a story about how one comes to have a CA of water, such as how someone's desire to quench their thirst is, under normal conditions, satisfied in part because of the role a prototype of water plays in guiding their recognition of water. But what of a CA of things like justice? Perhaps a thorough understanding of how the desires of political activists get satisfied would shed some light on the question, but this is conjecture.

Finally, while CSs are meant to underly all of our propositional attitudes, the CSs covered here and in the psychological literature are exclusively descriptive in nature, i.e. they represent how the thinker takes the world to be. There is little to no account of how CAs would acquire their identities in the context of directive representations. Considering that under teleosemantic approaches that focus on use such directive representations are more basic than

³⁸⁴ See Shea (2018) pg. 148-162, for discussion of these issues. Another possible response to worries about indeterminacy is to restrict what sorts of normal explanations are relevant. Millikan (1984), to take one example, holds that the content of a representation is determined by not just any normal explanation of the normal functioning of the mechanisms that make use of that representation, but rather the most "perspicuous" normal explanation available.

the descriptive representations they function in concert with, this is a major omission. However, this is I think more reflective of a general lack of theories of directive representations comparable to the well-known theories of descriptive conceptual representations in psychology. And as with indeterminacy, this is a problem that faces teleosemantic theories as a whole.

What I have argued for here is, again, not intended to be a well-worked out and well-supported theory of CS and CA semantics. What I have given is more of a gesture in the direction of a much more developed theory. It is a direction that I suspect will be more fruitful than its competitors. And I suspect it is best equipped to deliver what seem to be deep features of CSs, such as their adequacy being context-relative, their loose possession conditions, and their high degree of variability.

Conclusion

It is incredibly difficult to write anything new on concepts. Between the work of both psychologists and philosophers over the past hundred years (along with closely allied fields in the cognitive sciences more generally), there is more intellectual output than any one person could ever hope to take in and understand. I am reminded of a quote of dubious provenance often attributed to Kant, one of those that primarily traffics in social media feeds and countless quote aggregation websites. Despite its unlikely connection to Kant, I cannot help but find appealing its imagery, and find it comforting in moments of philosophical frustration:

Metaphysics is a dark ocean without shore or lighthouse, strewn with many a metaphysical wreck.

It is meant, I suspect, to capture that very real feeling of aimlessness and confusion that often comes with philosophical inquiry, where many times the questions you are investigating are as obscure as any of the potential answers you are seeking. Questions where you have not the slightest idea where to even begin. At times I felt this while researching and writing for this dissertation. But often I felt the problem was the opposite: rather than lacking guidance, I felt overwhelmed by its abundance.

The mind is now investigated by many fields, each with many competing and inconsistent paradigms. Whenever you start to grow confident that you have hit upon the right path, you come across a new group of thinkers. They have arguments as compelling as any of the ones you have grown sympathetic to. But if they are right, then everything you have ever thought or written about the mind is mistaken down to its foundations. It is as if, to draw on the metaphor from before, you are trying to navigate that dark ocean of uncertainty, hoping

that the one lighthouse you have chosen to guide you, that one lighthouse among hundreds of others, will not lead to you dash your ship upon the rocks. My ship, as of now, remains afloat. I can detect no leaks nor glint of light reflected from waves crashing against rocks in the night. I can only hope that it carries me safely to the shore.

I suspect that much, if not most, of what I have written here is not wholly original. Some of what I have developed is likely old ideas arrived at by new paths. Some is likely old ideas seen from a slightly different perspective than those who came upon them before me. Perhaps, if I am lucky, I have managed to add something novel to the immense stockpile of ideas regarding the mind. Despite worries of this sort, there are at least a few ideas present in this work that I think stand a greater chance at novelty than the others. And these, I think, warrant a brief consideration here, at the end of it all.

The first and most important of these is the relation between conceptual atoms and conceptual structures. It is not uncommon for traditional atomists to allow for the existence of conceptual structures like prototypes. Even Fodor allowed for as much, despite his famous hostility to many of the theories of concepts in psychology. They are, however, usually off to the side of concepts, serving as things such as conceptions, sustaining mechanisms, and other mental phenomena of lesser importance. Their role on this way of thinking is to assist the atomistic concepts in performing their own, more important functions, such as setting up and sustaining a content-determining relation between atom and world. Perhaps the operation of the conceptual structure serves as an enabling condition for the functioning of the atom as a part of a compositional conceptual system, this system perhaps filled with formulae of mentalese or expressions in a language of thought.

On my view, the relation between atoms and structures is much more intimate and intertwined. Neither conceptual atom nor structure is prior to nor more important than the other. Contentful conceptual structures can only exist given the existence of their component atoms, having referential content inherited by their cores and cognitive content inherited from their peripheries. Likewise, contentful atoms can only exist in the context of a conceptual structure of which they are a part. One cannot check an atom in isolation for the right mind-world relation to see what it represents. One must instead look at how the atom contributes to the functioning of the many conceptual structures of which it is a part. And understanding those structures requires attending to their constituent atoms and their many relations to one another.

The second idea is that of robust pluralism. This is a pluralism not just of coarse but of fine structured conceptual kinds, along with an acknowledgement and stressing of the immense variability present between such structures. This was, I think, the most important lesson I took from Millikan's work on concepts (and later unicepts/unitrackers), as whatever theory she may have been developing at the time, this aspect was always at the forefront. It is what convinced me that demanding identity of cognitive content would inevitably render publicity for concepts impossible. Likewise my account of concept individuation and the hierarchy of concept kinds owes much to Millikan's work on substances (and later clumps), along with her refreshingly uncompromising realism about everything from gothic cathedrals to waltzes. I have attempted, as best I can, to import her many insights into the framework for conceptual representations I have developed here.

Third, and finally, is the demand that conceptual atoms and structures be investigated in an uncompromisingly ecological manner. To understand how these representations work, you must look at what their possessors are trying to do with them, where they are trying to do

it, and what means they have available to them to achieve those ends. Representations that have seemed hopelessly inadequate or naïve to many philosophers can be remarkably well adapted given the right environments, user capacities, and ends. It is no fault of a thinker that they do not develop unnecessarily sophisticated cognitive tools for the form of life they live. It is not the fault of the toad that it tries to eat the BBs the psychologists shoot near it; eating small, black, moving objects would have worked to get food had the toad not been ripped from its supportive environment. Similarly it is no fault of mine when I eat an oatmeal raisin cookie thinking it to be chocolate chip. My cookie-detection mechanisms look only for small, black, cookie-embedded objects, and this strategy leads me astray only in a highly abnormal environment where raisin-filled cookies are served. In most cases, both the toad's and my own conceptual resources do a plenty good job getting us what we want and need.

I will end this dissertation by reflecting on what I take to be some of the main lessons of the reflections contained within. First, and perhaps most important, is the need to look at how conceptual structures are made use of by actual thinkers within realistic contexts. Too often in philosophy, discussions of concepts involve highly rarified and artificial examples, such as what kinds of inferences one must find compelling in order to grasp a concept of the various logical connectives, or what one's concepts would represent if transported to a mimic Earth where XYZ flowed through rivers and streams rather than H₂O. There is comparatively little discussion of by what representational means a young child recognizes their mother, or how it is that someone can so effortlessly bring to bear their knowledge of a recipe's ingredients to prepare a dish and in what structures such knowledge is encoded. And given a supportive enough environment, thinkers often require far less robust conceptual structures than I think has often been assumed.

Relatedly, considerations of conceptual variability point towards a thorough externalism about mental content for conceptual atoms and structures. Conceptual representations vary radically in their structures, both coarse and fine, and are often inaccurate in how they represent their referents. Internalist theories cannot accommodate this while still rendering concepts both public, stable, and possessable enough to serve as the ground for psychological explanations, both folk and scientific. Internalism, like molecularism, demands too much of thinkers.

I do not take the arguments I have given throughout this dissertation to be conclusive. Issues of cognitive architecture and taxonomization remain highly controversial, as do issues concerning intentionality and the extent to which cognitive science should adopt an ecological lens. Even issues as seemingly basic as the explanatory role for mental representation are fraught, with some denying even the existence and usefulness of mental content entirely. The ultimate test of any theory or framework in cognitive science is how explanatorily productive it is, and whether it smooths out the process of theorizing, both as individuals and in collaboration. Whether or not this is the case is not something that can be determined for any view a priori, including my own. My hope here at the end of it all is simply that something contained herein may be of use, or even just of interest, to someone out there. And I hope that it serves to bring me or someone who reads it just a little bit closer to that dark ocean's shores.

Works Cited

- Aristotle, & Barnes, J. (2014). *The complete works of Aristotle: The revised Oxford translation*. Princeton University Press.
- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, *13*(3), 263–308. [https://doi.org/10.1016/0010-0277\(83\)90012-4](https://doi.org/10.1016/0010-0277(83)90012-4)
- Ashby, F. G., & Rosedahl, L. (2017). A neural interpretation of exemplar theory. *Psychological Review*, *124*(4), 472–482. <https://doi.org/10.1037/rev0000064>
- Aydede, M. (1998). Fodor on concepts and Frege puzzles. *Pacific Philosophical Quarterly*, *79*(4), 289–294. <https://doi.org/10.1111/1468-0114.00063>
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*(1), 82–93. <https://doi.org/10.3758/BF03197629>
- Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227. <https://doi.org/10.3758/BF03196968>
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 629–654. <https://doi.org/10.1037/0278-7393.11.1-4.629>
- Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization* (pp. 101–140). Cambridge University Press.
- Barsalou, L. W. (1991). Deriving categories to achieve goals. In *Psychology of Learning and Motivation* (Vol. 27, pp. 1–64). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60120-6](https://doi.org/10.1016/S0079-7421(08)60120-6)
- Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. C. Collins, S. E. Gathercole, & M. A. Conway (Eds.), *Theories of memory* (pp. 29–101). Lawrence Erlbaum Associates. <https://philpapers.org/rec/BARFSA>
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. <https://doi.org/10.1017/S0140525X99002149>
- Barsalou, L. W., & Hale, C. R. (1992). Components of conceptual representation. From feature lists to recursive frames. In *Categories and Concepts: Theoretical Views and Inductive Data Analysis*. Academic Press. <https://philpapers.org/rec/BARCOG>
- Barsalou, L. W., Kyle Simmons, W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*(2), 84–91. [https://doi.org/10.1016/S1364-6613\(02\)00029-3](https://doi.org/10.1016/S1364-6613(02)00029-3)

- Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80*(3, Pt.2), 1–46. <https://doi.org/10.1037/h0027577>
- Block, N. (1998). Holism, mental and semantic. In E. Craig (Ed.), *The Routledge Encyclopedia of Philosophy*. Routledge.
- Block, N., & Campell, J. (1987). Functional role and truth conditions. *Aristotelian Society Supplementary Volume*, *61*(1), 157–184. <https://doi.org/10.1093/aristoteliansupp/61.1.157>
- Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition*, *60*(1), 1–29. [https://doi.org/10.1016/0010-0277\(95\)00699-0](https://doi.org/10.1016/0010-0277(95)00699-0)
- Boghossian, P. A. (1989). Content and self-knowledge: *Philosophical Topics*, *17*(1), 5–26. <https://doi.org/10.5840/philtopics198917110>
- Bowman, C. R., Iwashita, T., & Zeithamova, D. (2020). Tracking prototype and exemplar representations in the brain across learning. *eLife*, *9*, e59360. <https://doi.org/10.7554/eLife.59360>
- Boyd, R. (1999). Kinds, complexity, and multiple realization. *Philosophical Studies*, *95*(1–2), 67–98. <https://philpapers.org/rec/BOYKCA>
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 3–170). Lawrence Erlbaum Associates. <https://philpapers.org/rec/BRONCF>
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1986). *A study of thinking*. Transaction Books.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1999). The process of concept attainment. In E. Margolis & S. Laurence (Eds.), *Concepts: Core Readings* (pp. 101–123). MIT Press.
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, *4*, 73–121. <https://doi.org/10.1111/j.1475-4975.1979.tb00374.x>
- Burge, T. (1988). Individualism and self-knowledge: *Journal of Philosophy*, *85*(11), 649–663. <https://doi.org/10.5840/jphil1988851112>
- Camp, E. (2007). Thinking with maps. *Philosophical Perspectives*, *21*(1), 145–182. <https://doi.org/10.1111/j.1520-8583.2007.00124.x>
- Camp, E. (2009). Putting thoughts to work: Concepts, systematicity, and stimulus-independence. *Philosophy and Phenomenological Research*, *78*(2), 275–311. <https://doi.org/10.1111/j.1933-1592.2009.00245.x>
- Camp, E. (2015). Logical concepts and associative characterizations. In E. Margolis & S. Laurence (Eds.), *The Conceptual Mind* (pp. 591–622). The MIT Press. <https://doi.org/10.7551/mitpress/9383.003.0034>

- Carey, S. (1985). *Conceptual change in childhood*. MIT Press.
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. In E. Margolis & S. Laurence (Eds.), *The Conceptual Mind: New Directions in the Study of Concepts* (pp. 543–566). The MIT Press. <https://doi.org/10.7551/mitpress/9383.003.0031>
- Chomsky, N. (2000). *New horizons in the study of language and mind* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811937>
- Cohen, B., & Murphy, G. L. (1984). Models of concepts. *Cognitive Science*, 8(1), 27–58. https://doi.org/10.1207/s15516709cog0801_2
- Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, 72(20), 741. <https://doi.org/10.2307/2024640>
- Cummins, R. (1977). Programs in the explanation of behavior. *Philosophy of Science*, 44(2), 269–287. <https://doi.org/10.1086/288742>
- Cummins, R. (1983). *The nature of psychological explanation*. MIT Press.
- Daley, K. (2010). The structure of lexical concepts. *Philosophical Studies*, 150(3), 349–372. <https://doi.org/10.1007/s11098-009-9411-9>
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33(1–2), 25–62. [https://doi.org/10.1016/0010-0277\(89\)90005-X](https://doi.org/10.1016/0010-0277(89)90005-X)
- Dretske, F. (1981). *Knowledge and the flow of information*. MIT/Bradford Press.
- Dretske, F. (1986). Misrepresentation. In R. J. Bogdan (Ed.), *Belief: Form, content, and function* (pp. 17–36). Oxford.
- Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. MIT/Bradford.
- Dummett, M. (1993). *The seas of language*. Oxford University Press.
- Dupré, J. (1993). *The disorder of things: Metaphysical foundations for the disunity of science*. Harvard University Press.
- Edwards, K. (2009). What concepts do. *Synthese*, 170(2), 289–310. <https://doi.org/10.1007/s11229-009-9584-y>
- Edwards, K. (2011). Higher-level concepts and their heterogeneous implementations: A polemical review of Edouard Machery's *Doing Without Concepts*. *Philosophical Psychology*, 24(1), 119–133. <https://doi.org/10.1080/09515089.2010.544520>
- Estes, W. K. (1994). *Classification and Cognition*. Oxford University Press.

- Evans, C. S., Macedonia, J. M., & Marler, P. (1993). Effects of apparent size and speed on the response of chickens, *Gallus gallus*, to computer-generated simulations of aerial predators. *Animal Behaviour*, 46(1), 1–11. <https://doi.org/10.1006/anbe.1993.1156>
- Evans, V., & Green, M. (2006). *Cognitive linguistics: An introduction*. Lawrence Erlbaum Associates Publishers.
- Fodor, J. A. (1975). *The language of thought*. Crowell.
- Fodor, J. A. (1981). The present status of the innateness controversy. In J. A. Fodor (Ed.), *Representations: Philosophical essays on the foundation of cognitive science* (pp. 257–316). MIT Press. <https://philpapers.org/rec/FODTPS>
- Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. MIT Press.
- Fodor, J. A. (1990a). A Theory of Content I: The Problem. In *A theory of content and other essays* (pp. 51–87). MIT Press.
- Fodor, J. A. (1990b). A theory of content II: the theory. In *A theory of content and other essays* (pp. 89–136). MIT Press.
- Fodor, J. A. (1990c). Information and representation. In P. Hanson (Ed.), *Information, language, and cognition* (pp. 175–190). University of British Columbia Press.
- Fodor, J. A. (1995). Concepts: A potboiler. *Philosophical Issues*, 6, 1. <https://doi.org/10.2307/1523025>
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Clarendon Press ; Oxford University Press.
- Fodor, J. A. (2003). *Hume variations*. Clarendon Press ; Oxford University Press.
- Fodor, J. A. (2004a). Having concepts: A brief refutation of the twentieth century. *Mind and Language*, 19(1), 29–47. <https://doi.org/10.1111/j.1468-0017.2004.00245.x>
- Fodor, J. A. (2004b). Water's water everywhere [Review of Kripke: *Names, Necessity and Identity*, by C. Hughes]. *London Review of Books*, 26(20). <https://www.lrb.co.uk/the-paper/v26/n20/jerry-fodor/water-s-water-everywhere>
- Fodor, J. A. (2008). *LOT 2: The language of thought revisited*. Clarendon Press ; Oxford University Press.
- Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8(3), 263–367. [https://doi.org/10.1016/0010-0277\(80\)90008-6](https://doi.org/10.1016/0010-0277(80)90008-6)
- Frankish, K. (2010). Dual-process and dual-system theories of reasoning: Dual-process and dual-system theories of reasoning. *Philosophy Compass*, 5(10), 914–926. <https://doi.org/10.1111/j.1747-9991.2010.00330.x>

- Garrod, S., & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, 16(1), 77–90. [https://doi.org/10.1016/S0022-5371\(77\)80009-1](https://doi.org/10.1016/S0022-5371(77)80009-1)
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.
- Gelman, S. A., Coley, J. D., & Gottfried, G. M. (1994). Essentialist beliefs in children: The acquisition of concepts and theories. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind* (1st ed., pp. 341–366). Cambridge University Press. <https://doi.org/10.1017/CBO9780511752902.014>
- Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, 38(3), 213–244. [https://doi.org/10.1016/0010-0277\(91\)90007-Q](https://doi.org/10.1016/0010-0277(91)90007-Q)
- Gibson, M. (1996). Asymmetric dependencies, ideal conditions, and meaning. *Philosophical Psychology*, 9(2), 235–259. <https://doi.org/10.1080/09515089608573182>
- Glasgow, J. (2020). Conceptual revolution. In J. Glasgow, *Shifting Concepts* (pp. 148–166). Oxford University Press. <https://doi.org/10.1093/oso/9780198803331.003.0009>
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20(1), 1–19. <https://doi.org/10.1017/S0140525X97000010>
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66(3), 325–331. <https://doi.org/10.1111/j.2044-8295.1975.tb01468.x>
- Godfrey-Smith, P. (1996). *Complexity and the function of mind in nature*. Cambridge University Press.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. MIT Press.
- Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: What early learning tells us about the mind*. William Morrow & Co.
- Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, 37(5), 620–629. <https://doi.org/10.1037/0012-1649.37.5.620>
- Graham, P. J. (2012). Epistemic entitlement. *Noûs*, 46(3), 449–482. <https://doi.org/10.1111/j.1468-0068.2010.00815.x>
- Hampton, J. A. (1982). A demonstration of intransitivity in natural categories. *Cognition*, 12(2), 151–164. [https://doi.org/10.1016/0010-0277\(82\)90010-5](https://doi.org/10.1016/0010-0277(82)90010-5)
- Hampton, J. A. (1999). Conceptions of concepts. In B. Kokinov (Ed.), *Perspectives in Cognitive Science* (Vol. 4, pp. 27–38). Sofia: New Bulgarian University Press.

- Hampton, J. A. (2006). Concepts as prototypes. In *Psychology of Learning and Motivation* (Vol. 46, pp. 79–113). Elsevier. [https://doi.org/10.1016/S0079-7421\(06\)46003-5](https://doi.org/10.1016/S0079-7421(06)46003-5)
- Hampton, J. A. (2020). Investigating differences in people’s concept representations. In T. Marques & A. Wikforss (Eds.), *Shifting Concepts* (pp. 67–82). Oxford University Press. <https://doi.org/10.1093/oso/9780198803331.003.0005>
- Hampton, J. A., & Passanisi, A. (2016). When intensions do not map onto extensions: Individual differences in conceptualization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(4), 505–523. <https://doi.org/10.1037/xlm0000198>
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, 93(4), 411–428. <https://doi.org/10.1037/0033-295X.93.4.411>
- Hoffman, D. (1998). *Visual intelligence: How we create what we see*. W. W. Norton & Co.
- Hoffman, D. D. (2003). Does perception replicate the external world? *Behavioral and Brain Sciences*, 26(4), 415–416. <https://doi.org/10.1017/S0140525X03290099>
- Horowitz, A. (2017). Smelling themselves: Dogs investigate their own odours longer when modified in an “olfactory mirror” test. *Behavioural Processes*, 143, 17–24. <https://doi.org/10.1016/j.beproc.2017.08.001>
- Hull, C. L. (1920). Quantitative aspects of evolution of concepts: An experimental study. *Psychological Monographs*, 28(1), i–86. <https://doi.org/10.1037/h0093130>
- Jackendoff, R. (1983). *Semantics and cognition*. MIT Press.
- Jackendoff, R. (1990). *Semantic structures*. MIT Press.
- Jackendoff, R. (2011). 30. Conceptual semantics. In C. Maienborn, K. V. Heusinger, & P. Portner (Eds.), *Semantics*. De Gruyter Mouton. <https://doi.org/10.1515/9783110226614.688>
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. MIT Press.
- Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning and Verbal Behavior*, 23(2), 221–236. [https://doi.org/10.1016/S0022-5371\(84\)90148-8](https://doi.org/10.1016/S0022-5371(84)90148-8)
- Khalidi, M. A. (2013). *Natural categories and human kinds: Classification in the natural and social sciences*. Cambridge University Press.
- Khalidi, M. A. (2023). *Natural kinds* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009008655>
- Kintsch, W. (1974). *The representation of meaning in memory*. Lawrence Erlbaum Associates.
- Kitcher, P. (1984). Species. *Philosophy of Science*, 51(2), 308–333. <https://doi.org/10.1086/289182>

- Kitcher, P. (1993). *Inductive inference and its natural ground*. MIT Press.
- Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin*, 112(3), 500–526. <https://doi.org/10.1037/0033-2909.112.3.500>
- Kripke, S. A. (1980). *Naming and necessity: Lectures given to the princeton university philosophy colloquium*. Harvard University Press.
- Kwong, J. M. C. (2007). Is conceptual atomism a plausible theory of concepts? *The Southern Journal of Philosophy*, 45(3), 413–434. <https://doi.org/10.1111/j.2041-6962.2007.tb00058.x>
- Lalumera, E. (2010). Concepts are a functional kind. *Behavioral and Brain Sciences*, 33(2–3), 217–218. <https://doi.org/10.1017/S0140525X10000403>
- LaPorte, J. (1996). Chemical kind term reference and the discovery of essence. *Noûs*, 30(1), 112. <https://doi.org/10.2307/2216306>
- Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings* (pp. 3–81). MIT Press.
- Laurence, S., & Margolis, E. (2002). Radical concept nativism. *Cognition*, 86(1), 25–55. [https://doi.org/10.1016/S0010-0277\(02\)00127-0](https://doi.org/10.1016/S0010-0277(02)00127-0)
- Locke, J. (1689). *An essay concerning human understanding*. Oxford University.
- Machery, E. (2005). Concepts are not a natural kind. *Philosophy of Science*, 72(3), 444–467. <https://doi.org/10.1086/498473>
- Machery, E. (2009). *Doing without concepts*. Oxford University Press.
- Machery, E. (2010). Précis of *Doing without Concepts*. *Behavioral and Brain Sciences*, 33(2–3), 195–206. <https://doi.org/10.1017/S0140525X09991531>
- Machery, E. (2017). *Philosophy within its proper bounds* (First edition). Oxford University Press.
- Malt, B. C. (1994). Water is not h₂o. *Cognitive Psychology*, 27(1), 41–70. <https://doi.org/10.1006/cogp.1994.1011>
- Malt, B. C. (2010). Why we should do without concepts. *Mind & Language*, 25(5), 622–633. <https://doi.org/10.1111/j.1468-0017.2010.01405.x>
- Malt, B. C., Ameel, E., Gennari, S., Imai, M., Saji, N., & Majid, A. (2011). Do words reveal concepts? *Proceedings of the Annual Meeting of the Cognitive Science Society*, 33, 519–524.
- Margolis, E. (1998). How to acquire a concept. *Mind and Language*, 13(3), 347–369. <https://doi.org/10.1111/1468-0017.00081>

- Margolis, E., & Laurence, S. (2007). The ontology of concepts—Abstract objects or mental representations? *Nous*, 41(4), 561–593. <https://doi.org/10.1111/j.1468-0068.2007.00663.x>
- Margolis, E., & Laurence, S. (2013). In defense of nativism. *Philosophical Studies*, 165(2), 693–718. <https://doi.org/10.1007/s11098-012-9972-x>
- McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, 6(4), 462–472. <https://doi.org/10.3758/BF03197480>
- Mcgilvray, J. (1998). Meanings are syntactically individuated and found in the head. *Mind & Language*, 13(2), 225–280. <https://doi.org/10.1111/1468-0017.00076>
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–196). Cambridge University Press.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207–238. <https://doi.org/10.1037/0033-295X.85.3.207>
- Medin, D. L., & Schwanenflugel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory*, 7(5), 355–368. <https://doi.org/10.1037/0278-7393.7.5.355>
- Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, 35(1), 113–138. <https://doi.org/10.1146/annurev.ps.35.020184.000553>
- Meints, K., Plunkett, K., & Harris, P. L. (1999). When does and ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology*, 35(4), 1072–1078. <https://doi.org/10.1037/0012-1649.35.4.1072>
- Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society*, 7(3), 283–284. <https://doi.org/10.3758/BF03337190>
- Mervis, C. B., & Pani, J. R. (1980). Acquisition of basic object categories. *Cognitive Psychology*, 12(4), 496–522. [https://doi.org/10.1016/0010-0285\(80\)90018-3](https://doi.org/10.1016/0010-0285(80)90018-3)
- Michel, C. (2022). A hybrid account of concepts within the predictive processing paradigm. *Review of Philosophy and Psychology*. <https://doi.org/10.1007/s13164-022-00648-8>
- Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. MIT Press.
- Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science*, 56(2), 288–302. <https://doi.org/10.1086/289488>
- Millikan, R. G. (1993). *White queen psychology and other essays for alice*. MIT Press.
- Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs, and real kinds: More Mama, more milk, and more mouse. *Behavioral and Brain Sciences*, 21(1), 55–65. <https://doi.org/10.1017/S0140525X98000405>

- Millikan, R. G. (2000). *On clear and confused ideas: An essay about substance concepts*. Cambridge University Press.
- Millikan, R. G. (2005). *Language: A biological model*. Clarendon Press ; Oxford University Press.
- Millikan, R. G. (2017). *Beyond concepts: Unicepts, language, and natural information* (First edition). Oxford University Press.
- Mishler, B. D., & Brandon, R. N. (1987). Individuality, pluralism, and the phylogenetic species concept. *Biology & Philosophy*, 2(4), 397–414.
<https://doi.org/10.1007/BF00127698>
- Murphy, G. L. (1993). Theories and concept formation. In I. van Mechelen, J. A. Hampton, R. S. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 173–200). London, United Kingdom: Academic Press.
- Murphy, G. L. (2004). *The big book of concepts*. MIT Press.
- Murphy, G. L. (2020). How to make psychological generalizations when concepts differ: A case study of conceptual development. In T. Marques & A. Wikforss (Eds.), *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability* (pp. 34–50). Oxford University Press.
- Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1), 70–84.
<https://doi.org/10.1037/0278-7393.11.1.70>
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92(3), 289–316. <https://doi.org/10.1037/0033-295X.92.3.289>
- Neander, K. (2006). Content for cognitive science. In G. Macdonald & D. Papineau (Eds.), *Teleosemantics*. Oxford University Press.
- Neander, K. (2017). *A mark of the mental: A defence of informational teleosemantics*. MIT Press.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115(1), 39–57.
<https://doi.org/10.1037/0096-3445.115.1.39>
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9(1), 35–58. [https://doi.org/10.1016/0010-0277\(81\)90013-5](https://doi.org/10.1016/0010-0277(81)90013-5)
- Peacocke, C. (1989). What are concepts? *Midwest Studies in Philosophy*, 14, 1–28.
<https://doi.org/10.1111/j.1475-4975.1989.tb00180.x>
- Peacocke, C. (1991). The metaphysics of concepts. *Mind*, C(400), 525–546.
<https://doi.org/10.1093/mind/C.400.525>

- Peacocke, C. (1992). *A study of concepts*. MIT Press.
- Peacocke, C. (1993). How are a priori truths possible? *European Journal of Philosophy*, 1(2), 175–199. <https://doi.org/10.1111/j.1468-0378.1993.tb00031.x>
- Peacocke, C. (1996). Can possession conditions individuate concepts? *Philosophy and Phenomenological Research*, 56(2), 433. <https://doi.org/10.2307/2108535>
- Peacocke, C. (2004). Interrelations: Concepts, knowledge, reference and structure. *Mind and Language*, 19(1), 85–98. <https://doi.org/10.1111/j.1468-0017.2004.00249.x>
- Peacocke, C. (2009). Concepts and possession conditions. In B. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford University Press.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. The MIT Press.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77(3, Pt.1), 353–363. <https://doi.org/10.1037/h0025953>
- Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, 83(2, Pt.1), 304–308. <https://doi.org/10.1037/h0028558>
- Prakash, C., Stephens, K. D., Hoffman, D. D., Singh, M., & Fields, C. (2021). Fitness beats truth in the evolution of perception. *Acta Biotheoretica*, 69(3), 319–341. <https://doi.org/10.1007/s10441-020-09400-0>
- Prinz, J., & Clark, A. (2004). Putting concepts to work: Some thoughts for the twentyfirst century. *Mind and Language*, 19(1), 57–69. <https://doi.org/10.1111/j.1468-0017.2004.00247.x>
- Prinz, J. J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. The MIT Press. <https://doi.org/10.7551/mitpress/3169.001.0001>
- Putnam, H. (1975). The meaning of “meaning.” *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Rehder, B. (2003). Categorization as causal reasoning. *Cognitive Science*, 27(5), 709–748. https://doi.org/10.1207/s15516709cog2705_2
- Rey, G. (1983). Concepts and stereotypes. *Cognition*, 15(1–3), 237–262. [https://doi.org/10.1016/0010-0277\(83\)90044-6](https://doi.org/10.1016/0010-0277(83)90044-6)
- Rey, G. (1985). Concepts and conceptions: A reply to smith, medin and rips. *Cognition*, 19(3), 297–303. [https://doi.org/10.1016/0010-0277\(85\)90037-X](https://doi.org/10.1016/0010-0277(85)90037-X)
- Rey, G. (2009). Review of Edouard Machery, *Doing Without Concepts*. *Notre Dame Philosophical Reviews*, 2009(7).

- Rey, G. (2010). Concepts versus conceptions (Again). *Behavioral and Brain Sciences*, 33(2–3), 221–222. <https://doi.org/10.1017/S0140525X10000440>
- Rey, G. (2014). Analytic, a priori, false – And maybe non-conceptual. *European Journal of Analytic Philosophy*, 10(2), 85–110.
- Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 665–681. [https://doi.org/10.1016/S0022-5371\(75\)80055-7](https://doi.org/10.1016/S0022-5371(75)80055-7)
- Rips, L. J. (1995). The current status of research on concept combination. *Mind & Language*, 10(1–2), 72–104. <https://doi.org/10.1111/j.1468-0017.1995.tb00006.x>
- Rips, L. J. (1989). Similarity, typicality, and categorization. In A. Ortony & S. Vosniadou (Eds.), *Similarity and Analogical Reasoning* (pp. 21–59). Cambridge University Press. <https://doi.org/10.1017/CBO9780511529863.004>
- Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12(1), 1–20. [https://doi.org/10.1016/S0022-5371\(73\)80056-8](https://doi.org/10.1016/S0022-5371(73)80056-8)
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In *Cognitive Development and Acquisition of Language* (pp. 111–144). Elsevier. <https://doi.org/10.1016/B978-0-12-505850-6.50010-4>
- Rosch, E. (1978). Principles of categorization . In A. Collins & E. E. Smith (Eds.), *Readings in Cognitive Science: A Perspective From Psychology and Artificial Intelligence* (pp. 312–322). Morgan Kaufmann Publishers.
- Rosch, E. (1988). Principles of categorization. In *Readings in Cognitive Science* (pp. 312–322). Elsevier. <https://doi.org/10.1016/B978-1-4832-1446-7.50028-5>
- Rosch, E. (1999). Reclaiming concepts. *Journal of Consciousness Studies*, 6(11–12), 61–77.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605. [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, 2(4), 491–502. <https://doi.org/10.1037/0096-1523.2.4.491>
- Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology*, 15(3), 346–378. [https://doi.org/10.1016/0010-0285\(83\)90012-9](https://doi.org/10.1016/0010-0285(83)90012-9)
- Shea, N. (2018). *Representation in cognitive science* (First edition). Oxford University Press.
- Smith, E. E., Balzano, G. J., & Walker, G. J. (1978). Nominal, perceptual, and semantic codes in picture categorization. In J. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition*. Erlbaum Associates.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press.

- Smith, E. E., Medin, D. L., & Rips, L. J. (1984). A psychological approach to concepts: Comments on Rey's "Concepts and stereotypes." *Cognition*, *17*(3), 265–274. [https://doi.org/10.1016/0010-0277\(84\)90010-6](https://doi.org/10.1016/0010-0277(84)90010-6)
- Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, *8*(4), 337–361. https://doi.org/10.1207/s15516709cog0804_2
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*(3), 214–241. <https://doi.org/10.1037/h0036351>
- Solomon, K. O., Medin, D. L., & Lynch, E. (1999). Concepts do more than categorize. *Trends in Cognitive Sciences*, *3*(3), 99–105. [https://doi.org/10.1016/S1364-6613\(99\)01288-7](https://doi.org/10.1016/S1364-6613(99)01288-7)
- Sperber, D., & Wilson, D. (1998). The mapping between the mental and the public lexicon. In J. Boucher & P. Carruthers (Eds.), *Language and Thought: Interdisciplinary Themes* (pp. 184–200). Cambridge University Press. <https://doi.org/10.1017/CBO9780511597909.012>
- VandeWall, H. (2007). Why water is not h_2o , and other critiques of essentialist ontology from the philosophy of chemistry. *Philosophy of Science*, *74*(5), 906–919. <https://doi.org/10.1086/525632>
- Verheyen, S., & Storms, G. (2013). A mixture approach to vagueness and ambiguity. *PLoS ONE*, *8*(5), e63507. <https://doi.org/10.1371/journal.pone.0063507>
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., & Zilles, K. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage*, *14*(1), 170–181. <https://doi.org/10.1006/nimg.2001.0789>
- Wallis, C. (1995). Asymmetric dependence, representation, and cognitive science. *The Southern Journal of Philosophy*, *33*(3), 373–401. <https://doi.org/10.1111/j.2041-6962.1995.tb00751.x>
- Weisberg, M. (2006). Water is not h_2o . In D. Baird, E. Scerri, & L. McIntyre (Eds.), *Philosophy Of Chemistry: Synthesis of a New Discipline* (pp. 337–345). Springer Netherlands. https://doi.org/10.1007/1-4020-3261-7_18
- Weiskopf, D. A. (2009a). Atomism, pluralism, and conceptual content. *Philosophy and Phenomenological Research*, *79*(1), 131–163. <https://doi.org/10.1111/j.1933-1592.2009.00269.x>
- Weiskopf, D. A. (2009b). The plurality of concepts. *Synthese*, *169*(1), 145–173. <https://doi.org/10.1007/s11229-008-9340-8>
- Wilson, M. (2006). *Wandering significance: An essay on conceptual behavior*. Oxford University Press.

Wright, L. (1973). Functions. *The Philosophical Review*, 82(2), 139.
<https://doi.org/10.2307/2183766>

Zee, J., Storms, G., & Verheyen, S. (2014). Violations of the local independence assumption in categorization. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36). <https://escholarship.org/uc/item/1jj364dr>