9-2019

Some Non-Human Languages of Thought

Nicolas J. Porot

The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Comparative Psychology Commons, Philosophy of Mind Commons, and the Philosophy of Science Commons

Recommended Citation

https://academicworks.cuny.edu/gc_etds/3393

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
SOME NON-HUMAN LANGUAGES OF THOUGHT

by

NICOLAS POROT

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2019
Some Non-Human Languages of Thought

by

Nicolas Porot

This manuscript has been read and accepted for the Graduate Faculty in [program] in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

THE CITY UNIVERSITY OF NEW YORK
ABSTRACT

Some Non-Human Languages of Thought

by

Nicolas Porot

Advisor: Eric Mandelbaum

This dissertation asks: What might we learn if we take seriously the possibility that some non-human animals possess languages of thought (LoTs)? It looks at the ways in which this strategy can help us better understand the cognition and behavior of several non-human species. In doing so, it offers support, from disparate pieces of the phylogenetic tree, for an abductive argument for the presence of LoTs. Chapter One introduces this project. Chapter Two describes some axes along which LoTs might differ. It characterizes LoTs as collections of representations that may be combined to form others. It catalogues a non-exhaustive list of grammars that might permit such combination. These include grammars that allow for the formation of complex concepts, sentences, and complex sentences. It also explores the possibility that some LoTs might permit the syntactic combination of language-like representations with pictures, maps, analogue magnitudes, or vector representations. Chapter Three goes through all of the main arguments LoT Hypothesis. It argues that each fails to adequately support the presence of LoTs when applied to non-human animals (even if one accepts their application to human beings). Chapter Four focuses on arthropod mentality. It levies evidence, based on an underappreciated diagnostic tool—tests of multi-stability—for mental representation, and thus mentality, in Drosophila. It contrasts this case with a half dozen pieces of experimental evidence from bumblebees, paper wasps, and honeybees. These suggest that all three animals can form preference orders for their representations, and that honeybees possess a capacity to symbolically represent natural numbers, including zero, and at least up to six. In all cases,
explanations involving LoTs fare better than the most plausible alternatives, which rely mostly on associative learning mechanisms and analogue magnitude representations. Chapter 6 looks at recent evidence from chimpanzees, olive baboons, and an African grey parrot that suggest these species are competent with disjunctive syllogism. Since this competence requires inferences that rely on mental counterparts of disjunction and negation, they likely possess a LoT. It argues against alternative, Bayesian explanations of these results. And it considers some consequential upshots for our understanding logical inference in human development. Chapter Six presents the case for non-human LoTs on the strength of the evidence in Chapters Four and Five, and briefly describes possible future research in the study non-human LoTs.
for Jeanne Porot
ACKNOWLEDGEMENTS

As with most things, there are more people that deserve credit for this dissertation than I am going to be able to mention here. I realize it may seem trite to thank one’s advisor, but it would be a comical lapse of judgment not to start by thanking Eric Mandelbaum for the help he gave me in this process, and in the innumerable steps leading up to it. When I thought I couldn’t possibly ask him for more help or in a more annoying or inconvenient way, he somehow found a way to offer some I wouldn’t have dared ask for. I gladly accepted. Thanks also to the other members of the committee: Kate Ritchie, Liz Camp, Stephen Neale, and Michael Devitt for their help. I would like to thank Kate especially since she put up with multiple drafts of this, and provided extremely helpful feedback. I’d also like to acknowledge the Philosophy and Psychology departments at Yale, where I wrote the largest chunk of this dissertation as a visitor during the last year. Zoltan Szabo and Josh Knobe gave up time for conversations and comments on a paper that would become Chapters Three and Five in this dissertation. Those changed my thinking about this topic. Thanks also to Laurie Santos for tolerating my enthusiastic spectating in her canine cognition lab. There are others that helped long before all this. I’d not have made the starting gun in this process but for the CUNY graduate student community, and especially: Joseph Bendana, Tyler Wilbers, Jake Quilty-Dunn, and Henry Shevlin. Thanks to Michael Devitt for making me take writing seriously. Thanks to Moby for inspiring me to write this thing, and for keeping my feet warm.
## TABLE OF CONTENTS

### CHAPTER ONE: INTRODUCTION  
IX

### CHAPTER TWO: SOME LANGUAGES OF THOUGHT  
6  
Representations and Combinations  
8  
Combinatorial Minds  
17  
Recursion and Combination  
28

### CHAPTER THREE: A LACK OF EVIDENCE FOR NON-HUMAN LOTS  
47  
Languages of Thought and “Non-Linguistic” Animals  
47  
Some Obvious Ways to Tell if an Animal Has a Language of Thought, and Why They Don’t Work  
50

### CHAPTER FOUR: MINDS WITHOUT SPINES: MENTAL REPRESENTATION IN THE FRUIT FLY, AND ARTHROPOD LANGUAGES OF THOUGHT  
61  
Mental Representation in the Fruit Fly  
61  
Three Arthropod Languages of Thought  
77  
Number in the Honeybee LoT  
90  
Conclusion  
102

### CHAPTER FIVE: SOME EVIDENCE FOR DISJUNCTION AND NEGATION IN BABOONS, CHIMPANZEES, AND AN AFRICAN GREY PARROT  
104  
The Evidence  
104  
Upshots for the Development of Disjunctive Syllogistic Reasoning in Humans  
114  
What Was at Stake?  
120

### CHAPTER SIX: CONCLUSION  
123

### REFERENCES  
125
LIST OF TABLES AND FIGURES

TABLES

Table 1: Some Kinds of Combinatoriality for LoTs

FIGURES

Figure 1: Fruit flies were presented with an ambiguous motion stimulus. Reproduced from Toepfer et al. (2017).
Figure 2: Bumblebee pushing a black ball near the center of the platform. The bees were trained through social learning to place yellow, farther away balls (like the ones above) in the center of the platform.
Figure 3: Polistes wasps learned associations of colors (here: indigo) and shock when those colors appeared with others (here: violet).
Figure 4: Which of these clusters is larger in number?
Figure 5: Now, which of these clusters is larger in number?
Figure 6: Subitizing allows us to recognize small numerosities quickly.
Figure 7: Honeybees using signs in a maze to navigate to a sucrose reward. The bees learned to associate pairings of numerosities of elements (e.g. 2) and arbitrary symbol (e.g. “T”) with reward. Even when they only saw the items sequentially, they were able to exploit this association to locate the reward in a maze. Reproduced from Howard 2019a.
Figure 8: Honeybees learned to associate signs with three items on them with a sucrose reward. They selected cards with three items on them even when the authors varied the low-level features of the items on the card, and when they varied whether three was fewer or greater than the number of items on the alternative card. Reproduced from Bortot et al. 2019.
Figure 9: Honeybees learned to exploit “greater-than” and “less-than” relations for shapes to navigate to sucrose rewards they cannot see—even when the relevant relation is between one item and zero items. Reproduced from Howard et al. 2019b.
Figure 10: Infants look longer at outcomes that violate disjunctive syllogism than those that do not, suggesting they are surprised by this outcome.
Chapter One: Introduction

Everyone who has ever had a dog has wondered, at some point, what its thoughts were like. More generally, and in contrast with much philosophical theorizing, it seems obvious that a large number of non-human animal species have thoughts—as well as desires, intentions, perceptions, and other mental states. Yet, when we reflect on what those states might be like, and in particular, how they might be similar or dissimilar from our own, we realize that we really have very little sure footing from which to start. This dissertation is an attempt to answer that question, or part of it, at least, by starting with some relatively sure footing. It is a test run, an experiment using a familiar theoretical tool, to see how far it can get us. It asks: What might we find if we take seriously the plausibility of non-human Languages of Thought (henceforth, LoTs)?

In the service of that end, this dissertation has two goals. The first is fairly armchair-philosophical: to consider various ways that a LoT, and thus a LoT-equipped mind, mind be. I will argue that there is quite a bit more flexibility in what counts as a LoT than might first appear. The second is strictly empirical: to survey results in comparative psychology from a wide range of species—from fruit flies to chimpanzees—to see how far different versions of LoT-style explanation can get us. It offers arguments that versions of that theoretical posit offer the best explanations for each of them, further strengthening the explanatory case for LoTs generally. And it catalogues various kinds of LoT that might exist in non-human minds (animal or otherwise).

---

1 I will also speak of a LoT, or LoTs, rather than the LoT. The reason for this will become apparent in Chapter Two, where I discuss ways in which languages of thought might differ.
LoTs were introduced as a posit of the Language of Thought Hypothesis (LoTH). The LoTH is an account of how minds represent, and of how minds transition between representational states. Minds represent by means of mental symbols, realized in the brain, that can be combined in the manner of formal language symbols. Minds transition between states by performing computations over those symbols, in the manner of a formal language. (I will have more to say about these conditions in later chapters). Proponents of the LoTH, in particular Jerry Fodor (1975, 2008; Fodor & Pylyshyn 1988), often rejoin that LoTs are the primary or only means by which human minds might achieve the feats they do. The LoTH, and in particular this rejoinder, have been a significant object of contention in the philosophy of mind for over four decades now. Because, in this project, we are concerned with non-human minds, where the relevant question is whether there is any LoT to speak of, we can drop this rejoinder, though we need not. Either way, from here on, I will often speak of LoTs, rather than of the LoTH.

Whether any mind (human or otherwise) possesses a LoT is consequential. For one thing, if such languages exist at all, they surely form a psychological kind. Merely in virtue of having minds that are representational and combinatorial in the way required by LoTs, any minds with a LoT will resemble one another.² For another thing, whether or not one agrees that there is a human LoT, it is plausible that LoTs could underpin many of the psychological competences its proponents have said it underpins for human psychology. A mind with a sufficiently syntactically powerful LoT is capable of natural language acquisition, production, and understanding; concept learning; logical inference; mathematical reasoning; the bearing of propositional attitudes; and perception, at least. Thus, evidence that an animal possesses a LoT is also evidence that it possess some or all of these

² LoTs, like the neurons that typically realize them, might emerge more than once in evolutionary history (neurons emerged at least twice: once in our lineage, and once in that of Ctenophores). This view is compatible with their being a psychological kind. The grammar of a LoT, as Chapters 4 and 5 of this dissertation bear witness, is defined functionally.
competences—even if it does not realize all or any of these competences because of performance constraints. This has upshots for how humans relate to other species. If humans possess a LoT, its presence in non-human animals would provide a kind of cognitive continuity between us and those animals. (And, if not, the origin of cognitive life is all the more mysterious). LoTs can even work as a kind of diagnostic tool for recognizing animals whose cognition resembles ours.

And, as I will argue in Chapters Four and Five, there is good reason to suppose that at least some non-human minds possess a LoT. New methods in comparative psychology are emerging rapidly, as those working in that field have moved away from old behaviorist assumptions about animal learning and behavior. It is not uncommon today to read, in the discussion sections of comparative psychology journal articles, sincere uses of such mentalist expressions as “concept,” (Howard et. al. 2019a) , “inference” (Pepperberg et al. 2018), or “metacognition” (Rosati & Santos 2016). And, of course, one often finds things when one begins to look for them. In recent years, empirical data in support of behaviors in non-human animals that are strongly at odds with those old behaviorist assumptions have poured in. One does not go from rejecting behaviorism to advocating for LoTs, of course. But a surprising amount of these results are congenial to explanation in terms of LoTs. And some of the evidence would require impressive feats of parameterizing to explain without them. Moreover, the evidence cited ranges over distant clades of Animalia. Some of it comes from familiar faces in discussions of non-human minds: great apes, monkeys, and birds. But, in a marvel of experimental evidence overcoming intuition, other evidence comes from a more surprising source: arthropods.

---

3 Which competences it is evidence for will depend on the syntax of the LoT in question. I take this up in Chapter 2.
There is a slow-moving inference to the best explanation afoot in this strategy, of course. By accumulating enough snapshots of the animal world that are congenial to explanation in this way, we garner some traction for actually explaining them this way. And so, to the degree that one is willing to step back away from the details of explaining the cognitive abilities of individual species, a broader picture begins to emerge. The view that we applied so as to get some sure footing will turn out to be supported by the evidence we inspect, strengthening the assumption we took on.

Chapter Two is the part of the dissertation that handles the armchair-philosophical element of this project. It gives a minimal characterization of LoT's and taxonomizes ways LoT's might differ from one another. These differences concern not just the kinds of representations that a LoT might include in its lexicon, but also different ways those representation might be combined. In addition to providing us with a framework for thinking about minds that are different from ours, this taxonomical project has the added benefit of helping us get clear on the central features of LoT’s—which are, on their own, an historically important theoretical posit.

Chapter Three considers several arguments for the Language of Thought Hypothesis, some traditional and some novel, as applied to non-human animals. I argue that they all fail. Even if one finds such arguments compelling in the case of human beings, they lend at best middling support to the claim that any other species possess a LoT. This is especially evident in the most familiar arguments, which rely heavily on natural language. It is thus an open question whether any non-human animal does in fact possess a LoT.

Then, in Chapters Four and Five I address the empirical question of which species might actually possess LoT's, if we are looking for them. Chapter Three uses four arthropod species as case
studies. First, I consider the case of the fruit fly, a species which, surprisingly enough, appears to visually represent at least one feature of its environment. I appeal to some recent empirical evidence and argue, at some length, against the most plausible non-representational interpretation of the empirical evidence in support of this claim. And I advocate for an underappreciated tool for detecting mental representation in non-linguistic species: tests of multi-stable perception. I then move on to consider honeybees, bumblebees and paper wasps, three species which appear to be able to rely on syntactic combination of representations in the service of foraging. I again argue against competing explanations of these capacities, which rely on association and analogue magnitude representations. I close by considering some evidence that the honeybee LoT possesses discrete symbols for number, a hallmark of re-combinable symbolic representation.

In Chapter Five, I consider empirical evidence for representation of logical connectives in non-human animals. I survey empirical evidence from chimpanzees, baboons, and an African grey parrot named Griffin as examples. Research on these animals suggests they are capable of a form of disjunctive syllogism, suggesting the possession of disjunction-like and negation-like representations in their repertoires. I conclude by considering the state of our knowledge about the lexicon of their LoTs and some upshots for the range of competences that might be underpinned by such representations.
Chapter Two: Some Languages of Thought

For all that has been said about the LoTH, surprisingly little has been said about LoTs themselves. It is possible, and seems likely, that if LoTs do occur in this world, some differ from others. This chapter explores axes along which such languages might differ. Of course, there are lots of ways that things can differ, some of which are ways that things do not in fact differ. Some of the LoTs considered here might not feature in any extant mind. To be clear, this chapter explores possible differences.

There are at least three reasons to care about this kind of project. First, it is inherently interesting to explore the logical space LoTs occupy. They have held an outsized role in the philosophy of mind for a few dozen years, even when that role was punching bag. Whether or not one considers the LoTH viable, it should be of interest to know what its central posit is, exactly.

Second, and more practically, considering different features of a LoT can help us to understand what it would take for a system to possess one. It could help us understand which creatures—and indeed, which computers, extraterrestrials, or ant colonies—share this feature of their minds. This is consequential: Traditionally, it has been assumed that LoTs, if they exist, play a widespread role in cognition. They are meant to partly explain how cognitive abilities such as thought, inference, perception, reasoning, the language faculty, and the bearing of propositional attitudes, are psychologically possible. One need not think they play such a widespread role. But to the degree that they do, one can expect that animals (as well as machines, Martians, ant colonies, and so on.) with LoTs will share such features of their psychologies.
Finally, this kind of project can help us understand how LoTs might differ from one another. In looking at other minds than our own, we do not want to infer the presence of some features of LoTs merely because we have detected the presence of others. Learning about the ways that LoTs differ can help avoid this kind of error. If I order mole off the menu of a Mexican restaurant, I am to blame if I am surprised when the meal arrives, and it is not covered in a chocolate-based sauce. It may be covered, not in mole negro, but in Oaxacan red or yellow mole. These share some of the many ingredients of mole negro, but do not contain chocolate. This kind of assumption can be avoided by just learning about how moles differ. I will argue that, as with mole, there is much variety in LoTs. And so, we are at fault if we move, unjustifiably, from identifying one feature we know to be indicative of LoTs to inferring the presence of other features commonly associated with them. In this way, it helps to know which features of a LoT are dissociable from one another. In Chapters Four and Five, I describe some empirical evidence of non-human LoTs that we must approach with these distinctions in mind. The evidence will leave open what kind of LoT each species possesses.

In Section I, I consider shared features of LoTs: They are collections of mental representations, and they are combinatorial. In Section II, I consider different ways a collection of mental representations might be combinatorial. They might differ in virtue of the kind of combinatoriality at play, or in virtue of the ways in which a language might be combinatorial without being recursive. Finally, in Section III I argue, against convention, that many sorts of mental representations—such as pictures and maps—might be combined syntactically with linguistic symbols, or with language-like symbols and non-linguistic representations. Consequently, some possible LoTs contain representations that lack predicate-argument structure.
Representations and Combinations

We can start with shared features of LoT’s. First off, perhaps the most minimal claim one can make about a LoT is that it is a collection of mental representations. This is fairly uncontroversial. That said, there are also some reasonable arguments for it. First, the LoTH, by which the concept of a LoT was introduced, is a species of the computational-representational theory of mind (CRTM). That view holds that the mind is made up of mental representations, and that its movements are best explained as computations over those representations. LoTs, on this view, are the representations a mind computes over, or at least a significant number of them. So understood, it is trivial that a LoT is a collection of mental representations.

Then again, one need not tether a theoretical posit to its etiology. So maybe LoTs are not collections of mental representations? But LoTs, whatever one thinks of them, are languages. And representing is an important function of language. To take an example from natural language, when I say

(1) The 7-11 is three blocks down on your right

---

4 Even Jerry Fodor (1975, Ch. 5, and especially 2008, Ch 1.) was happy to admit that some mental representations, for example those invoked in mental rotation and other imagery tasks, were not part of the LoT, though, so far as I can tell, he did not go so far as to claim they might interact with the propositional structures of the putative human LoT. Others, especially Pylyshyn (2003) were more ambitious. They maintain that to the degree that mental imagery even really calls for explanation, rather than for explaining away, such an explanation can be provided by a LoT. In any case, all parties here agree that the LoT is a set of mental representations over which computation occurs.
my words really represent some place and a certain trajectory from where you are. This seems a stronger reason for accepting that languages of thought are representations.

Of course, I have not explained what I mean by ‘mental representation’. I take it that nothing I will say throughout the rest of this chapter, or indeed the rest of this dissertation, will hinge on a particular theory of representation. I load ‘representation’ with as little as possible: as something that can be true or false—or, at a minimum, accurate or inaccurate. It is whatever makes it the case that my beliefs can be true or false, and my perceptions accurate or inaccurate. I will have a bit more to say about this when I discuss, in Chapter Three, minds that mentally represent, but nonetheless lack LoTs. In any case, that LoTS are mental representations seems not really up for debate. Whatever a LoT might be, it is supposed to play a role in explaining some feature of psychology.

Another claim about LoTs that is fairly uncontroversial is that they are combinatorial. The representations in the language, and their meanings, can be combined to form others. What it means, more specifically, for some language to be combinatorial differs depending on whether one is considering the syntax of the language or its semantics. I will talk about each in turn, before turning to some varieties of syntactic combination.

On the syntactic end, combinatoriality just amounts to putting representations together to make new ones. For example, suppose my dog has concepts, and those concepts are representations in his LoT. By syntactic combinatoriality, if my dog has the concepts BONE and TASTY, he can put them together to form TASTY BONE. Note that, though this example resembles a modified noun
phrase in English, where ‘tasty’ modifies ‘bone’, the LoT string need not be parsed into a modifier and a noun. A LoT might have a syntax that lacks modification, or lacks predicate-argument structure altogether. All the matters here is that there are two representational types (TASTY and BONE), and that there is some way to put together tokens of them that result in a token of a third kind (TASTY BONE).

The putting together of representations must happen in accordance with rules of the LoT in question.⁵ These rules need not be explicitly represented by the system that stores the LoT. And of course, not all rules can be explicitly represented. Even if my dog psychologically represents all of the syntactic rules of his LoT, the rules that govern their application must be represented only implicitly—or else the rules that govern those rules must be, and so forth…on pain of regress. One reason for supposing that no syntactic rules of the LoT are explicitly represented is simply that for this to be possible, there would need to be some metalanguage in which the rules of the LoT are composed. The rules in that metalanguage could not themselves be represented, of course, unless there was another metalanguage, and so forth… again, on pain of regress. Nothing rules out the possibility of explicitly represented metalinguistic rules, but that route seems at best unappealing, if it can be avoided. Instead, the syntactic rules might be mere descriptions of regularities in the way the system combines representations.

---

⁵ For there to be rules, there must be combinations of representations that are well-formed according to such rules. But there need not be representations that are not well-formed. The reason is that one can imagine a language in which any string may be combined with any other. It seems likely that only languages with extremely simple combination rules or extremely limited numbers of representations could work this way. Still, such languages might exist.

Still, this is a fairly minimalist specification of what those syntactic rules might be like, and a more fleshed out account may rule in a much more precise characterizations of what it takes for a syntactic rule to obtain. One possibility, explored primarily in the human LoT but applicable more widely, is that the rules are representations that are built into cognitive architecture: they are representations that need to be tokened by a particular system in any of its transformations (Quilty-Dunn & Mandelbaum 2018)
One the semantic end, the question is different. Are certain meanings determined by other meanings, and the way those other meanings are combined? Or are meanings determined at least partly in some other way? On the traditional construal of the LoTH, the semantics of a LoT perfectly mirrors the syntax of that LoT. So, to each syntactic part, there corresponds one and only one semantic part, which is the meaning of that syntactic part; to each syntactic combination, there corresponds some isomorphic semantic combination (presumably, a combination of meanings), which determines the meaning of the resulting string. The result of this view is that the meaning of each well-formed string in the LoT is determined (completely) by the syntactic structure of the string, and the meanings of the constituents of that string. So, for example, if TASTY and BONE are atomic (not composed) symbols then the meaning of TASTY BONE is determined by the meanings of TASTY and BONE and the way they are combined.

One need not tie the semantics of a LoT to its syntax in this way. Semantic properties of sentences in a LoT could be determined by other facts. For example, the meanings of strings might derive from their functional role in the cognitive system: how they contribute to inferences, perceptual judgments, and the like. So long as that role is not determined (solely) by the meanings of the constituent concepts and the way they are combined, a LoT could have a semantics that is at least partly determined by conceptual role, even though it has a combinatorial syntax. This is an odd view, since it is unclear exactly what the syntax of the LoT would be contributing to the overall cognitive life of the system (or exactly how we would measure it). But, at least to the degree that conceptual role semantics itself is not incoherent or otherwise hopeless, such a view is not incoherent, either. We are interested in whether a mind could function by means of such a language, not whether human minds (or any other) do. The same goes for any non-compositional view of
meaning one might have. And there are also other, spookier violations of the traditionally assumed relationship between an LoT’s syntax and its semantics. We can imagine a Frankenstein OR: FROR. Uses of it can be expressed, in English, by adding “and pigs” to the argument of the second sentence that it conjoins. So, when my dog thinks THAT IS A TASTY BONE FROR THIS COUCH IS A GOOD PLACE FOR A NAP, he is having a thought that can be expressed in English with “That is a tasty bone or this couch and pigs are a good place for a nap”. Other spooky cases are easy to come up with. And, as far as I can tell, nothing we have said so far about LoTs precludes them.

What does it really mean to say that the semantics of FROR, any such spooky connective, diverges from its syntax? If an English sentence with ‘or’ is the best expression of a particular LoT sentence, why not think that the LoT string it expresses contains OR? The point here is that a LoT that contains OR could just as well be such that it has spooky operators, for all we know. We are not interested in whether the putative human LoT (or, for that matter, any actual language) contains FROR. We are interested in whether some representational system might.

We can now look at the ways that the rules of syntactic combination might vary in a LoT. This question is broad. It amounts to asking how the rules of any combinatorial grammar might vary. A satisfying discussion of that would go beyond the ken of this dissertation. And, if there is no upper bound on that ways that grammars might differ, such a discussion might simply be impossible. In the rest of this section, I consider some straightforward variations on combination rules. I will mostly focus on examples that are likely to come in handy later. In interpreting some

---

6 This kind of connective could not obviously combine with sentences with dummy arguments, like ‘It’s raining’. So, to the degree that any LoT might have them, it would be an unhelpful syntactic rule for a LoT to possess. But imagine a Frankenstein variation on negation, FRANKENNOT. Expressions of it in English add “and pigs are highly social creatures” to whatever sentence it is conjoined with: “It’s not raining and pigs are highly social creatures.”
empirical evidence I present in Chapters Four and Five, the distinctions I raise here will help to avoid attributing too complex a LoT to the species under consideration.

There are at least three sorts of syntactic combinatoriality a LoT might display. First, there is what we can call *sub-sentential combinatoriality*. This occurs when representations combine to form complex, sub-sentential representations. For example, maybe my dog possesses a LoT. I am fairly confident he possesses the concepts BONE and TASTY. If his LoT has sub-sentential combinatoriality, he will be able to put those concepts together to form the complex concept TASTY BONE. As I watch a high fly ball arc back down to earth, my visual system tokens representations of roundness, of whiteness, and of downward motion. Plausibly, these features are combined by my visual system, forming a complex representation of a round white downward moving object. If that representation is part of my LoT, it would look something like: ROUND WHITE DOWNWARD MOVING OBJECT.

In both of the cases I just mentioned, the resulting representations are complex, since they have multiple constituents, but they are also sub-sentential, in the sense that they do not involve a complete predicate-argument structure. Of course, my dog, or I, could fail to have the working memory capacity needed to perform such combinations. Or he (or I) might not possess some other psychological capacity required for combination. Or, that capacity might be temporary impaired, for example because we ingested something toxic. In such cases, we would fail to form complex

---

7 There is a metaphysical question about how these concepts might combine. In virtue of what do TASTY and BONE form TASTY BONE? I see this question as fairly outside the ken of this discussion. It is not really under dispute that concepts do combine. In fact, that concepts can be combined is one claim that virtually all realists about concepts agree on. Suffice it to say that whatever makes it the case that concepts can combine will, by definition, be a feature of any language of thought (unless—for some reason—this feature of concepts should turn out to be incompatible with their being mental representations!). That said, there is at least one minimal condition worth pointing to: It is safe to say that whatever concatenation is, it cannot be mere association (for a discussion, Mandelbaum 2015).
The second kind of combinatoriality is *sentential combinatoriality*. This is the ability to combine concepts to form sentences, as in

(2) THAT IS A TASTY BONE.

Syntactically, this is very similar to sub-sentential combinatoriality.

How different are sub-sentential and sentential combinatoriality, really? This depends on what distinguishes sentences from other complex mental representations. One might think that the syntactic combination that is at play in forming complex concepts is not just very similar to, but *exactly the same as* the syntactic combination at play in forming sentences (e.g., Pietrowski 2011). And so it is silly to distinguish types of combinatoriality on these grounds. One reason to think this would be if one individuates sentences not by any of their syntactic features, but only by their semantic ones. Sentences, but not complex concepts, have sentential contents. I can form the belief that *it is going to be hot today*, but I really cannot form the belief that *hot day*. And perhaps that is all

---

8 Some can deny this. Dispositionalists about belief, such as Schwitzgebel (2002), offer one example. Since they think believing is a matter of acting in certain ways, they need not characterize belief as a relation to sentences, or to the contents of sentences. Of course, the LoTH loses much of its zing if LoTs do not figure centrally in explaining belief, and so dispositionalism is not *naturally* paired with LoTs, and is silent on mental representation more generally. Still, a dispositionalist about belief might still think that mental representations might explain other sorts of mental state, such as perceptual ones. Such a person should reject the distinction between sub-sentential and sentential combinatoriality. But, as Quilty-Dunn & Mandelbaum (2018) stress, such views fail to explain how beliefs can be truth apt. If one can believe that *hot day*, then there is only a tenuous relationship between truth-aptness and believing, while belief would seem *paradigmatically* truth-apt.
one needs to tell sentences from complex sub-sentential strings in a LoT—even if they turn out to be identical syntactically. Given this, it would be a virtue of any language of thought if, as Fodor assumed (1975), its syntax and semantics mirror one another, so that this worry does not really arise. But given only what has been said up to now, we cannot be sure whether there is any difference—syntactic or semantic—between these two types of combination.

Finally, there is supra-sentential combinatoriality. This consists in combinations of whole sentences with other representations. Sentential connectives offer one kind of example. If my dog possesses a conceptual counterpart of OR, he may form the complex LoT sentence

(3) THAT IS A TASTY BONE OR THIS COUCH IS A GOOD SPOT FOR A NAP.

<table>
<thead>
<tr>
<th>Sub-sentential only</th>
<th>Sentential only</th>
<th>Supra-sentential only</th>
<th>Sub-sentential + sentential</th>
<th>Sub-sentential + sentential + supra-sentential</th>
<th>Sentential + supra</th>
</tr>
</thead>
<tbody>
<tr>
<td>THAT BONE</td>
<td>WFF</td>
<td>*</td>
<td>*</td>
<td>WFF</td>
<td>WFF</td>
</tr>
<tr>
<td>THAT IS A BONE</td>
<td>*</td>
<td>WFF</td>
<td>*</td>
<td>WFF</td>
<td>WFF</td>
</tr>
<tr>
<td>THAT IS A BONE OR I AM HUNGRY</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>WFF</td>
<td>WFF</td>
</tr>
<tr>
<td>THAT IS A TASTY BONE OR THIS COUCH IS A GOOD SPOT FOR A NAP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>WFF</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 1: Some Kinds of Combinatoriality for LoTs

9 In fact, this feature actually has another virtue. It extends Alan Turing’s project of building rational machines. The tight connection between syntax and semantics is a defining feature of machine and programming languages, and of formal languages generally. So it would be very nifty if it were so for a language that in fact explains our own rational cognition—and equal parts disheartening if that language were not so.
This is two LoT sentences, combined with a connective, to form a third. A dominant view in logic, and a popular one for language, is that connectives may be combined with whole sentences. In the case of (classical) propositional logic, the connectives can only be connected with sentences. A LoT with sentential combination might be like propositional logic in this way. The important thing is that the kind of combination in question does not operate on sub-sentential parts of a LoT.

I conclude by discussing a few borderline cases of supra-sentential combinatoriality. These are cases of combinations of representations in a LoT that lacks predicate-argument structure. A LoT might possess individual symbols for sentences: THAT-COUCH-IS-A-GOOD-SPOT-FOR-A-NAP rather than THAT IS A GOOD SPOT FOR A NAP. By means of such symbols and OR, it could form complex representations, such as THAT-IS-A-TASTY-BONE OR THAT-COUCH-IS-A-GOOD-SPOT-FOR-A-NAP. Similarly, a mind might combine representations that are not often thought of as members of LoTs, such as pictures, maps, analogue magnitude representations, or vectors (section IV contains a longer discussion of these formats, but for now, we can just focus on the fact that they lack predicate-argument form). One can imagine, for example, a “disjunction” of picture-like mental representations.

A mind might lack sub-sentential or sentential combinatoriality but still possess one or more these strange connectives. Such a mind could not form LoT sentences with predicate-argument structure, and so it seems they would not possess supra-sentential combinatoriality, either. But what kind of combinatoriality do they possess? One response, the response that I will take, is to modulate supra-sentential combinatoriality slightly. Make it include combinations of one-word sentences, pictures, maps, and the like, with sentential connectives. One reason for doing this is that such representations might do many of the things sentences do in propositional calculus. Let ‘Merkel’
stand in for a picture of the German chancellor, writing at her desk, and ‘Macron’ stand in for a picture of the French President (where these are not names, but just shorthand for the pictures themselves). And let OR* and NOT* be connectives that can combine with pictures. A LoT might implement inferences of the form: Merkel OR* Macron; NOT* Merkel; Macron. There will be wrinkles to iron out later, in section IV: What might such connectives be like? Are they truth functional? How they might be different from OR and NOT? But it bears stressing how similar such connectives really are to OR and NOT. Another response is to propose another sort of combinatoriality to accommodate these connectives. One might find them too far removed from the connectives that operate over sentences. On that view, any mind that possessed such connectives would be combinatorial, even if it lacked all three of the kinds of combination I have described in this section. Little hinges on this taxonomical choice. All that will matter, later on, is that such connectives might exist in a mind of some kind, and that they bear important similarities to those of classical logic.

Combinatorial Minds

The three sorts of combination I have described, sub-sentential, sentential, and supra-sentential, tend to cluster together in discussions of combinatoriality, as well as in discussions of linguistic compositionality. This is understandable, since natural language tends to respect them, and, it would appear, so does human thought\textsuperscript{10}. And much discussion of LoTs has focused on human

\textsuperscript{10} Cf. Johnson 2004 for an interesting objection to this sort of inference. He maintains that minds that are not systematic might nonetheless express thoughts in a language that is systematic, handling a “tool that they could never fully make use of.” In the same article, he also levies evidence against the systematicity of language and questions the introspective obviousness of systematicity of thought, further undermining the most familiar forms of support for systematicity of thought.
minds—the evidence for which typically relies on human language. But a cognitive system might possess a competence for some of these kinds of combinatoriality, and not others. It is worth considering, in some detail, the representational limits and strengths of minds that are “incompletely” combinatorial in various ways.

To this end, I will consider what minds might be like that possess only one of each of these kinds of combination. This should not be taken to suggest these kinds of combination are exclusive of one another. In fact, it is not even likely such forms of combinatoriality typically occur independently of one another. There may be interactions between different kinds of combinatoriality. Perhaps a mind with two forms of combinatoriality possesses psychological capacities that two minds, each with one form of combinatoriality, jointly lack. Nonetheless, this seems as good an means as any of demonstrating just what each of these forms of combination might contribute to a mind. I'll call a mind that is capable only of sub-sentential combination a sub-sentential mind, a mind that is capable only of sentential combination a sentential mind, and a mind that is capable only of supra-sentential combination a supra-sentential mind.

Consider a sub-sentential mind. All it can do is form complex (sub-sentential) representations, like TASTY BONE. It cannot form LoT sentences, like THAT LOOKS TASTY. And of course, it cannot combine sentences or similar representations operators, like sentential connectives, that combine with sentences or similar representations. Despite its limitations, such a mind might actually perceive, recognize, and categorize a vast—potentially infinite—number of objects and states of affairs. That is because constraints on combinatoriality do not place limits on the number of representations such a mind might store in memory, or the size of the complex representations it might form by combination. With a large representational repertoire, this sort of mind could
demonstrate impressively environment-sensitive behavior. It could learn (or innately possess) a similarly vast number of associations between representational states and behaviors.

There are also a fairly large number of kinds of sub-sentential mind. Concepts and other representations might vary in the number of arguments that they can combine with to form larger representations. For example, imagine a concept DELIVERER(X,Y), which might be expressed in English as ‘deliverer of x to y’. It combines with two other representations to form a complex concept. (Though it has argument-places, this should not be thought of as a predicate: The result of the combination is a concept, not a sentence). Sub-sentential LoTs might vary with regard to the number of arguments a representation can take to form a larger representation, as well as with regard to the upper or lower bound of the adicities of those representations.

Sub-sentential minds can form object representations: They can represent objects as such. Commonly, in cognitive psychology, object representation is considered through the lens of object perception: How do people perceive objects as such (or do they do so at all)? But the phenomenon is more general, of course, since one can think about things just as well as one can perceive them. A sub-sentential mind can achieve object perception by at least three means. The first is mental demonstrative reference. Mental demonstratives are psychological counterparts of natural language demonstratives, such as English ‘that’. Zenon Pylyshyn (2009) and Tyler Burge (2010) have argued, independently, for versions of the view that mental demonstratives explain object perception. Because demonstratives function to refer to individuals, a representational state that includes a demonstrative element, such as THAT, can manage to refer to individuals. Such a view might be
implemented in a LoT\textsuperscript{11}: The language would need only possess a demonstrative, such as THAT. This view of object representation has been introduced in perception, but it can be extended to other capacities of the mind, including thought. The capacity for demonstrative reference allows for object representation tout court, not just object perception. And a sub-sentential mind that possesses a demonstrative symbol in its repertoire could achieve object representation in thought. A sub-sentential mind could even combine its THAT-symbol with other representations to form complex demonstratives, such as THAT BONE.

A second means of object representation that might be achieved by sub-sentential minds is reference by description. Traditional proponents of mental reference to particulars by description, such as Quine (1960), have limited the capacity primarily to linguistically competent adult humans, who have acquired natural language quantifiers. A trove of empirical work suggesting that there is object perception both in infants and many non-human animals makes the traditional version of this view implausible (The evidence may be familiar by now, but in any case, Ch. 1 of (Carey 2011) offers an overview). Still, such a view could be adapted to a LoT framework, without any assumptions about natural language. On that view, representation of objects is secured by definite descriptions in the LoT\textsuperscript{12}. Whether a sub-sentential mind is capable of mentally representing particulars in this way will depend on whether definite descriptions in it can occur alone as syntactically well-formed strings. On some views (Russell 1905, Neale 1990) they can appear meaningfully only in the context

\textsuperscript{11}Pylyshyn, but not Burge, advocates for the use of demonstratives in a LoT. Burge rejects the representational theory of mind, and so, extension, LoTs.

\textsuperscript{12}There are plenty of well-rehearsed objections to description theories for natural language, many of which are fatal to the simplest versions of them. But none that I am aware of are charges of incoherence. They all charge that description theories make the wrong predictions about expressions in natural language. But we are not concerned with evidence from natural language here. We care about how a mind might achieve reference to objects. So far as I can tell, there is no reason to suppose a mind could not function this way.
of a sentence. The latter route is barred for sub-sentential minds because they cannot form sentences. But the former route is available.

A final account of object representation, and one that is popular in the study of vision, is the “attentional feature-map” view (Treisman & Gelade 1980, Treisman 1998). On views of this kind, perceptual object representation occurs when maps of visual features are bound by focal attention, unifying features at locations into object representations. Such representations are unified by the shared locations of various features. What makes the representation of an object is the fact that features of the same objects (shape, and color, for example) share locations.

Feature maps could be encoded in a LoT. One means of achieving this is by means of a conjunction of sentences that predicate particular features (e.g. RED31) to particular locations (e.g. x1,y1). This route is of course not available to the sub-sentential mind, which lacks sentences and connectives. But a sub-sentential mind might achieve feature representation by means of one-to-many combination rules. Consider a four-place predicate: LOCATION(a,b,c,d). It assigns each of its arguments to a location in space. For example, in 2 x 2 space, LOCATION(BLUE42,RED3,GREEN12,YELLOW5) represents the following color grid:

x1y1: BLUE42
x1y2: RED3
x2y1: GREEN12
x2y2: YELLOW5
With a very large number of argument places, such a representation could be used to form very large, detailed feature maps. This is computationally cumbersome, but possible. Other predicates might also take arguments of different kinds, such as shape. And it might output not to individual locations in the perceptual field, but to larger areas. Finally, some predicates might take maps of different features (shape, color, etc) as inputs and output a bound representation of features at locations.

A sub-sentential mind is, surprisingly, capable of a minimal kind of logical thought. One might think that logical thought requires inferences of the sort we are familiar with from propositional or predicate logic (or else, something very close to them). In this case, the answer is clearly: No. Conventionally, the quantifiers of predicate logic figure only in full sentences. And the sentential connectives of propositional logic are conjoined only with full sentences. A sub-sentential combinatorial mind would have no need for LoT counterparts of logical quantifiers or connectives: Any combination it could form with them would be syntactically not well-formed. But other logical systems offer some hope for the sub-sentential mind. There is, for example, a particular tradition in the history of logic, stretching back to Aristotle, that allows for direct negation of predicates, rather than of sentences. In a LoT, this would mean, e.g., that NOT LIGHT is the negation of LIGHT. (Interestingly, there is evidence that some natural languages really do possess a form of constituent negation (Horn). So this may actually be a feature of the human LoT, if there is one). So there is a possible inference pattern exemplified by:

P1. NOT NOT F

C. F
This is a relatively uninteresting inference—if it is one. We would not want to say that any mind, in virtue of this capacity, is competent with rational thought. And moreover, there are deep and well-known problems with predicate negation. But a mind might be structured in this way. Such a mind would be capable of transitions in thought that adhere to rules. Those rules belong to something that historians of logic would recognize as logic. Nonetheless, such minds would unable to perform the bulk of rational inferences.

Now, we can turn to the strengths and limits of sentential minds: minds with sentential combinatoriality, but not sub-sentential or supra-sentential combinatoriality. Such minds can combine symbols to form sentences, but their representational repertoires are limited: They store no complex sub-sentential predicates. Such minds store atomic names and predicates in memory, but no complex representations. Such a mind could never think about tasty bones as such. Nonetheless, it could use simple representations, or whole sentences in its language, to perceive, recognize, categorize, and the like. A system of this kind might also be able to form associations between sentences and specific behaviors. With a large enough repertoire of representations in memory, there would be a similarly vast, potentially infinite, number of sentences such a mind could compose; consequently, it would be behaviorally very sensitive to its environment, much like a mind with only sub-sentential combinatoriality might be.

In addition, such a mind might possess a limited capacity for quantification. For example, if the syntax of its LoT resembles that of first-order predicate logic, it will be capable of universal and existential quantification. In predicate logic, quantifiers can be nodes in syntactic trees. So they do not need to be combined with predicates to form complex expressions. For example, consider the formula \( \exists x (\text{BONE}x) \). This has the quantifier, '\( \exists x \)', as its top node.
Thus, a mind with sentential but not sub-sentential combinatoriality might form sentences that include quantificatiers. It could form states syntactically similar to $\forall x (\text{TASTY}x)$, or $\exists x (\text{BONE}x)$. This is a large increase in the expressive power of any LoT. It allows for thoughts about generalities, such as the thought that everything is tasty, as well as thoughts about particulars, such as the thought that something is a bone.\(^{13}\)

Curiously, though, without representations to play a role similar to those of sentential connectives, such a mind would be incapable of nearly all of the inference patterns familiar from first-order logic. That is because for nearly all of those inference patterns, one needs sentential connectives. For example, ‘All dogs are mortal’ is treated as $\forall x (\text{DOG}x \supset \text{MORTAL}x)$ in predicate logic. Some limited inference patterns may be available to it; for example, existential generalization:

\[\exists \text{P. THAT IS A BONE}\]

\(^{13}\)Interestingly, however, this kind of mind cannot form representations that are syntactically similar to the quantifiers of natural language. In natural languages, quantifiers are typically parsed as constituents of determiner or noun phrases. These phrases are daughter nodes of the sentence or of another phrase or clause. For example, consider this (simplified) parsing of ‘All bones are tasty’:

\[
[\text{DP} [\text{DET} [\text{All}] [\text{N}\ [\text{bones}]] [\text{VP} [\text{V}\ [\text{are}]] [\text{AD}\ [\text{tasty}]]]]
\]

Here, the important thing is just that the quantifier ‘All’ is a constituent of the DP, which is a constituent of the sentence. A mind without sub-sentential combinatoriality cannot form such a representation. It is incapable of combining quantifiers with predicates to form complex constituents of sentences. Another way to say this is that it cannot form the complex representation \text{ALL BONES}.
C. SOMETHING IS A BONE

This is a capacity for a paradigmatic but fairly uninteresting form of deductive inference. On its own, it falls short of qualifying a mind as competent with logical thought. Nonetheless, it is a valid deductive inference, and so is available to the sub-sentential mind.

Sub-sentential minds may be capable of some forms object representation. First, they can represent individuals by means of “bare” demonstratives. They can put together sentences of the form THAT IS AN F. They can also form definite descriptions. For example, suppose a LoT contains the unique symbol DAUPHIN, which refers to heirs to throne of France. One could think sentence-sized thoughts including the description THE DAUPHIN by means of sentential combination alone: THE DAUPHIN IS LAZY. By contrast, no mind without sub-sentential combinatoriality could think thoughts including complex descriptions, such as THE KING OF FRANCE. That description includes a complex predicate, KING OF FRANCE. Finally, sentential combination, like sub-sentential combination, does not allow a mind to represent individuals by means of feature maps. A sentential mind could form feature maps of perceptual features at locations by means of large numbers of sentences (e.g. LOCATION X1,Y12 IS RED2. LOCATION X19,Y5 IS BLUE54). But they lack sentential connectives in their representational repertoires to conjoin those sentences. So it is unclear how such sentences could be strung together to form a single representation. (Again, an exception must be made if the mind in question contains cognitive maps, or pictorial representations).

A short note about quantifiers in a LoT: These need not possess all the logical properties of those we learn in first-order logic. Maybe no minds, including ours, really use such representations,

---

14 Note that this is different from representing particulars by means of complex demonstratives, like THAT F.
rather than approximate them. There are a wealth of alternative quantifier-representations a mind might possess. For example, instead of the familiar ALL, a mind might possess the quantifier PAUL: a quantifier that selects for everything, except when it takes CHILDREN in its scope, in which case it does not select for Paul. So, if a Martian thinks PAUL THE PLANETS IN THIS SOLAR SYSTEM ARE BORING, that might be expressed in English with “All the planets in this solar system are boring.” And if it thinks PAUL THE CHILDREN WENT FOR TEA AT THE GRAND MOSQUE OF PARIS, that might be expressed as: “All the children, except for Paul, went for tea at the Grand Mosque of Paris.” That is a strange example, of course, and there is no reason to think any creature might benefit from having such a syntactic rule in its LoT. But there is also no reason to think other, more adaptive variations on traditional quantifiers might not appear in the representational repertoires of some LoTs—including ours. And one could certainly program a robot with just such a quantifier in its mental lexicon.

There is one more, slightly more tendentious potential difference between minds with only sentential combinatoriality, and minds with only sub-sentential combinatoriality. On the traditional formulation of the LoTH, forming sentences in the LoT is part of what allows a mind to bear attitudes to contents: We have the propositional attitudes we do because we bear certain functional relations to sentences in our minds, and those sentences have contents. If we can bear the right relation to the sentence, we can take the corresponding attitude to its content. One need not accept such a picture of the nature of propositional attitudes to suppose that a cognitive system forms LoT sentences. One might think that LoTs are useful for such abilities as perception or categorization, but that activities such as belief fixation occur—for some reason—via a different process. Or one might have some alternative account of how beliefs relate to LoTs. But to the extent that one does accept a view of this kind, it is possible to say that a mind that can form LoT sentences meets at
least one necessary condition of believing, wishing, hoping, desiring, and the like. And a mind that cannot do so, does not.\(^\text{15}\)

We can now return to the third form of syntactic combination, *supra-sentential* combination. These, to remind the reader, are minds that can concatenate entire sentences with other representations. These include, paradigmatically, sentential connectives, like OR. But they could include any other elements of language that combine with whole sentences, rather than with sub-sentential constituents (I discuss an example below). If the kinds of borderline cases discussed at the beginning of these chapter are possible—combinations of one-word sentences, or of non-linguistic representations—LoT syntax might allow supra-sentential combinatoriality, but lack sub-sentential or sentential combinatoriality. Any mind without supra-sentential combinatoriality would lack the ability to concatenate sentences with logical connectives. It would be unable to perform deductive inferences of the kind familiar from propositional logic.

Since a mind with supra-sentential combination might possess conceptual counterparts to the logical connectives, it might also execute inferences of propositional calculus. These inferences need not take place over representations that are themselves built up from parts. Rather, they could take place over representations with sentential contents, but no syntactic constituents. So a sentential mind could execute them. Consider the following example. My dog might have a *SQUIRRELS-CLIMB-FAST* symbol in his LoT. So, he might think that squirrels climb fast by means of just that one symbol. If, in addition, he has an OR-like representation and a *MY-WAYWARD-MASTER-HAS-

---

\(^{15}\) As noted, though, it is not the LoT sentences that one bears attitudes to, but the sentences’ contents. So it would suffice if a concept, or any mental representation at all, had a sentential content. One’s mind could then bear the appropriate functional relation to that representation. It seems bizarre to say that a concept might have sentential content, but the objection is vastly more plausible for pictures, maps, and other sorts of mental representation. Perhaps a hyperrealist portrait of Simone Weil has a sentential content: It predicates large number of features to Simone Weil. I return to this question at the end of the chapter.
RETURNED representation in his repertoire, he can think SQUIRRELS-CLIMB-FAST OR MY-WAYWARD-MASTER-HAS-RETURNED —by means of just three symbols\(^\text{16}\). Similar considerations could apply to any operator that takes sentences as inputs. One can imagine, for example, tense and aspect markers that can be applied to whole sentences. For example: PAST: THAT-IS-A-TASTY-BONE, (roughly, in English: “That \textit{was} a tasty bone”), or PERFECT: THAT-IS-A-TASTY-BONE (“That has been a tasty bone”).\(^\text{17}\)

Recursion and Combination

\begin{quote}
LoTs are defined as combinatorial collections of mental representations. However, the rules that govern combination in a language need not be recursively defined. A mind might possess combinatoriality, though its transformations do not adhere to any recursively defined rules. A traditional focus on recursion and the related notion of productivity in discussions of language risks ignoring a large number of non-recursive combinatorial minds.
\end{quote}

First, consider the difference between combinatoriality and recursion. Combinatoriality, when considered in the context of language, is just the capacity to combine representations, in accordance with some syntactic rule or rules. Recursion is a structure that allows, by means of repeated applications of a syntactic rule, to combine representations to form representations of arbitrary

\(^{16}\) Or just two symbols: combination with logical connectives, like the execution of inferences, might be built into one’s cognitive architecture.

\(^{17}\) Typically, in natural languages, such markers are not actually combined with whole sentences, but with verb phrases in those sentences. This would not be enabled by sentential combinatoriality alone.
length. It is not hard to see how the two come apart. A language might place constraints on the number of iterated applications of a syntactic rule, so that while combinations are possible, those combinations cannot be iterated—or cannot be iterated an indefinite number of times. For example, a LoT might possess an OR-like representation that does not iterate. So, while such a mind could form A OR B and B OR C, it could not form A OR B OR C (even when one brackets any worries about performance constraints, such as working memory).

Before, moving on, a word about that last example, the uniterable OR-like representation. It might seem like a recursive rule is still needed to characterize the combinations of this representational type with others. This might be a rule similar to the one that allows for combinations of English sentences with ‘or’. It would simply contain an extra condition, which does not allow for application of the rule to disjunctions. But while such a rule could generate uniterable OR, a language could also have uniterable OR as primitive. In that language, a recursive rule for multiple iterations of OR would have to be defined by some addition to the syntactic rules of the language, or some other rule altogether. So far as I can tell, nothing tells against this possibility. Non-iterable OR is a possible LoT connective\(^\text{18}\). Similar examples can be imagined for all kinds of recursive rules.

This distinction matters when considering possible LoTs. If a collection of representations is a LoT in virtue of being combinatorial, then a LoT might possess combinatorial, but not recursive, rules across the board: It might lack any recursive rules. We have little armchair grip on the syntax

\[^{18}\text{This kind of example might seem arcane. But it will be especially useful to think about later, in empirical contexts, where non-human animals have only shown evidence for one or two iterations of OR. It will also prove useful in section IV. considering disjunctions of representations that lack predicate-argument structure. For example, imagine a form of disjunction, OR\(^*\), that only combines with pairs of pictures: PICTURE1 OR\(^*\) PICTURE2. OR\(^*\) could not be recursive, since it would result in a linguistic representation that could not be embedded in other iterations of OR. Thanks to Kate Ritchie for this latter point.}\]
of particular LoTs. This includes our own LoT, if we have one. But we are especially ignorant of the syntax of LoTs for animals and other cognitive systems that lack natural language (see Ch. 3, sec. II). So we have little if any way of knowing which expressions in a LoT, in particular for non-human animals, allow iterability.

The distinction between combination and recursion also shows a weakness in arguments from productivity in establishing the presence of a LoT: They grossly undersell the number of potential LoTs. It is often thought that recursive rules of combination are the best explanation for the productivity of language and thought. The familiar argument tends to run: *We can produce, understand, perceive, or think a boundless number of expressions, scenes, or thoughts. Our capacity to store mental representations is limited. So our mental representations must not be boundless in number. But a recursive structure is the best explanation for the emergence of the boundlessness from boundedness.* Now, plainly, recursion does lead to productivity. And so if productivity is the starting point of one’s discussion—as it usually is when considering the cognitive capacities and natural languages of humans—it is reasonable to conclude that the system under consideration is governed by recursive rules, rather than by rules that are merely combinatorial. But reliance on productivity risks greatly underselling how cheap combinatoriality really is: Many (possible) LoTs are not productive. This heightens the risk of failing to detect LoTs where they occur in the natural world.

Before closing, I add a terminological remark. So far, I have mostly avoided using the word “compositionality”, and in instead used “combinatoriality” to refer to the putting together of mental representations. This is particularly salient here, since arguments from productivity are often taken to support the compositionality of language and thought. I use ‘combinatoriality’ here in the interest of clarity. As Zoltan Szabo (2012) notes, ‘compositionality’ is closely associated with discussions of
natural language meaning, and so its use can be misleading in discussions of mental representation. Since a LoT might differ in important ways from natural languages, this association can load the dice somewhat in discussions of mental representation. They can lead people to make assumptions about LoTs that make them look more like natural language. For example, Szabo points out that linguistic expressions, but not concepts (read: mental representations), require interpretation. And if mental representations do not need interpretation, as Szabo thinks\(^{19}\), then talk of the compositionality of meaning of mental representations is simply superfluous. Another possible confusion is that since natural language appears to respect recursively defined rules, saying that LoTs are “compositional” might easily be taken to require such recursion. The familiar definition of compositionality—that the meaning of expressions is determined by the meanings of their parts and how they are combined—says nothing of whether there might be syntactic upper limits on those combinations. Still, speaking of “combination”, rather than “composition”, avoids this possible confusion.

To sum up, we have now gone over a non-exhaustive list of some different ways that LoTs might combine representations. They might combine them to form complex predicates, or sentences, or complex sentences. These combinatorial rules are independent of one another: A mind could possess some, but not others. Moreover, combinatorial minds might fail to be recursive or productive.

1. Cosmopolitan Languages of Thought

---

\(^{19}\) Of course, I have just argued in sec. II that this syntax-semantics mirroring is not a necessary feature of LoTs. But it is still suggestive of a difference between LoTs and natural language. Such mirroring is not a feature of any natural language I am aware of. This leaves open that it might not be a feature any natural language could possess. Meanwhile, it seems to be a perfectly normal feature to suppose at least some LoTs possess, since it reduces the possibility of ambiguity in strings of the LoT.
What kinds of mental representation can be combined according to the rules of a LoT?

What sorts of representation does a LoT contain? In this section, I explore the controversial idea that some LoTs might be cosmopolitan. They might include representations of more than one kind. I argue that a LoT might allow for syntactic combinations of representations of two or more formats and that this could motivate the view that some LoTs contain representations of those formats. I conclude the section by considering some objections to the view.

First we can say a little bit about what different kinds of representation are, and what it would mean for more than one to be part of a LoT. Psychologists and philosophers have proposed a range of kinds, or formats, of mental representation. Differently formatted representations encode information differently. Some representations have a predicate-argument structure similar to that of natural language. Concepts are often taken as an example. Such representations may allow for what I called earlier sub-sentential and sentential combination. Others do not have such a predicate-argument structure. Some represent states of affairs in the way that pictures do (Burge 2010; Block, Chapter 3, unpublished manuscript, Ch. 3; Carey 2011; Beck 2015; Quilty-Dunn 2019); others represent places, relations, or trajectories in a map-like way (Gould 1986, Rescorla 2009abc, Camp 2007); still others represent quantities as analogue magnitudes (Gallistel & Gelman 2000, Beck 2019), or word-meanings as vector representations20. It might seem, on the face of it, that only the language-like representations—those with a predicate-argument structure—should figure in a LoT. Yet as far as the the characterization we have been relying on is concerned, the question appears to be open. In line with tradition on this topic, I have not specified more about LoTs than that they are combinatorial, and that they are collections of mental representations. This says little of what the

---

20 Abstract relations may also be represented in a diagram; though formal treatments of diagrams do exist (Tufte 2012; Greenberg and Giardino, eds., 2015), to my knowledge diagrams have not been explored as a format of mental representation.
format of these representations might be like. It only (implicitly) constrains the list of possible formats to those that permit combination.

For each of the representational formats I have mentioned, there are many who doubt they play an explanatory role in human or non-human animal cognition. Some have insisted that the mind really is just made up of representations with a predicate-argument form (Pylyshyn 2002, 2003) or picture-like representations (Kosslyn 2006); others have delimited the range of specific formats to pieces of cognitive architecture—for example, limiting picture-like representations to perception (Block, unpublished manuscript, Beck 2016, Burge 2010, and Carey 2011). So am I on shaky ground when I imagine that any or all of them might be included in a LoT? Here, it is important to recall we are considering the limits of combinatorial representational systems generally, extant or not. So if a representational format is such that some mind (any mind) might store instances of it, that format should be on the table for consideration.

If the mind is cosmopolitan, being populated with different sorts of mental representation, it would be of certain value for such representations to be able to interface. The syntactic structure of a sentence offers a perfectly appealing way for this to occur (see below). Very likely, the mind is in fact cosmopolitan to some degree, as many, even Fodor himself, have stressed (Kosslyn 1978; Burge 2010; Gould 1986; Rescorla 2009a, Camp 2007, 2009; Fodor 1975, 2008; Beck 2015, 2019; Carey 2011; Block ms, Quilty-Dunn 2019). The utility of an interface of differently formatted representations is some reason to think a sophisticated cognitive system will interface those representations syntactically. This might seem a weak motivation, but it is also fairly widespread: Plausibly, it is part of the reason many find the idea of a language-like syntax for mental representations attractive in the first place.
Of course, this argument cuts both ways. There may be some a degree of integration possible with other representational formats, too. And if there is a benefit to interfacing in one direction, there is a benefit to interfacing in the other. Maps are an example: They can contain linguistic elements, such as YOU ARE HERE, that interface with spatial elements of the map, such as a distribution of green indicated a forest. For example, this might indicate that whoever is reading the map is in the middle of a particular forest. They can also integrate pictures, so that, e.g., a picture of the Mona Lisa in a map of the Louvre represents the Mona Lisa as being in a particular part of the Louvre. One can, similarly, imagine pictures that include maps, diagrams that include maps or pictures, and so on. These kinds of cases, in particular linguistic labels integrated into maps, have been explored elsewhere (Casati & Varzi 1999, Rescorla 2009a). Interestingly, however, the other sort of integration, of non-linguistic elements in mental states with that have a sentence-like structure, is underexplored (cf. Johnson 2015, Papineau 2002).

There is at least one plausible way that pictures, maps, magnitudes, or vectors might be integrated into a LoT sentence. Earlier, I pointed out that representations of those kinds might be the constituents of supra-sentential combination—or of a different kind of combination that is very similar to supra-sentential combination. They might be combined syntactically with conceptual counterparts of logical constants, such as OR. Alternatively, they might combine with slightly different connectives, such as OR*. The non-linguistic constituents of such combinations do not have an internal predicate-argument structure. Nonetheless, the resulting string does many of the things a sentence does in a supra-sentential combinatory LoT: It figures in inferences of a sort familiar from propositional calculus. If a LoT is a collection of representations, and this string is a sentence in the LoT, then that is some reason to think its constituents are part of the LoT. This
includes the pictures, maps, or other representations. Even if it is not a sentence, it bears important similarities to sentences: It is a complex representation, and it could well play a cognitive role similar to that of a sentence (for example, by figuring in inferences). This is a somewhat weaker, but still appealing, reason for thinking its constituents are members of a LoT.

The claim on the table is: If a mental representation could, in accordance with the syntactic rules of some LoT, be combined with another to form a complex representation, it is part of a LoT. So, for example, it suffices for a pictorial representation to count as part of my dog’s LoT if the syntax of his LoT allows that a particular picture could be conjoined with another picture (or some other representation) to form a complex representation in his mind. That could be so, even if that particular picture never figures in any such computation.

This rules out of LoTs some combinations we might otherwise expect to be ruled out. Suppose I learn to roll my eyes each and every time I token NEW YORK NICKS TRADE DEAL. The eye-rolling motor command would go together, in a sense, with that concept. But we should not want to say that there is some new, complex representation formed by the combination of that concept and the eye-rolling motor command (even if the command is made up of representations from my LoT). Likely, the concept is merely associated with the eye rolling command. If which representations figure into the LoT is determined by the syntactic rules of the language, there is a principled reason for distinguishing this case from other kinds of combination.

There are also some plausible limits to what representations might be included in the LoT. Reflection on two examples will be helpful here. Suppose that I have a LoT and that it includes complex language-like strings, such as I AM TURNED AROUND IN DOWNTOWN BROOKLYN. Suppose,
further that I have a cognitive map of Downtown Brooklyn. In the first case, we can imagine that the syntax of my LoT does not allow for combinations of cognitive maps with other representations. It seems plain here that my cognitive map of Downtown Brooklyn is not part of my LoT. Now, suppose that the syntax of my LoT is such that it permits such combinations. But my cognitive map of Downtown Brooklyn is stored in some part of my mind that makes it impossible for me to combine it with LoT strings—for example because it is stored in memory in an informationally encapsulated module. It seems at least plausible that in this case, unlike in the other case, the cognitive map might figure in the LoT, even though it never, in fact, combines with other representations. This is a prima facie reason for thinking that it is the syntactic rules of a LoT that determine which representations ought to count as part of it. The mere fact that the syntax of one’s LoT could have been otherwise is not reason for reconsidering which representations are part of one’s LoT. Whether some picture, map, or even sentence, is part of one’s LoT will depend on what kinds of combination are permitted by the syntactic rules of one’s LoT. I turn now to some objections and replies.

**Objection:** I have not specified the means by which pictures, maps, and other representations manage to interface with language-like representations. I cannot just stipulate that anything can be part of a disjunction, for example. That is a substantive claim.

**Reply:** This is a serious problem the study of LoTs must take up. But it is not an objection to the view that sentences in a LoT might include pictures, maps, magnitudes, or vectors. That is
because there is no adequate explanation of how sentences interface with connectives syntactically.

And so the problem is not really particular to the format of the representations in question.\textsuperscript{21}

It helps little to turn to traditional construals of the LoT here. On the traditional picture, due to Fodor, combination is possible because LoTs are innate, and the minds that possess them are “built to use” them (1975, pp. 67-8). Thus, in making an analogy with LoTs and machine languages, Fodor notes: “formulae can be paired directly with the computationally relevant states of the machine in such fashion that the operations the machine performs respect semantic constraints on formulae in the machine code” (ibid.). Taking this analogy at face value, this means the geometric properties of the realizing bit of brain (or silicon, or whatever) make some combinations, and thus some transformations, possible in a LoT. Since, further, Fodor assumes there is a neat mirroring of the properties of a LoT’s physical realizer, the syntax of that LoT, and its semantics, this claim serves a trifold job. It explains how not just the realizer, but also the syntactic and semantic constituents, can be combined. But as I pointed out earlier, the syntax of a LoT string might diverge in some ways from its semantics. And moreover, Fodor’s claim about machines being “built to use” representations is a bit mysterious—hardly less mysterious than simply asserting that representations can combine. This may well be a case in which there is little or nothing more to explain. But even if so, there is nothing about this idea rules out that a mind might be “built to use” pictures, maps, and the like in a way appropriate for combination with linguistic representations in a LoT.

\textsuperscript{21} In fact, a more specific version of the question, of how different \textit{kinds of mental state} might interface, is also not particular to me. At least one form of the phenomenal concepts picture, due to David Papineau (2002) must deal with it. That view calls for the interface of phenomenal states with other constituents of thought, in states of the form \textit{IT FEELS LIKE THAT}, where \textit{THAT} directly contributes a conscious qualitative state to the thought. Such views must explain how that interface is possible.
Syntactic rules govern combinations of representations. And a grammar might be such that a picture may play some of the same syntactic roles as a sentence. Any syntactic rules will specify which combinations are well-formed. Even some language-language combinations will typically turn out to be not well-formed. So a syntax might be such that some combinations involving non-linguistic representations are well-formed, and others not. Or it might be such that any combinations of linguistic and non-linguistic representations are not well-formed. But ultimately, much of the task of deciding which of these possibilities is actual will be an empirical question for any particular LoT.

It might seem that I am missing the point here. Perhaps there is a problem regarding how linguistic constituents interface syntactically. But even if there is, that does not preclude that there is another interface problem, over and above the humdrum one, that pictures, maps and other representations face. But articulating just what that other interface problem—the one that sentences, but not other sorts of representation, can address—is quite difficult to do. I think that is because there is none. There are things that natural language does that pictures, maps, and other types of representation don’t do. A sentence can be nominalized; a map can’t. A sentence can become a clause a picture (most likely) can’t. And so on. But these syntactic differences do not show that pictures, maps, and other representations could not play a limited number of the syntactic roles of sentences with predicate-argument structure; it merely shows that do not play all of them.

Of course, so far I have been talking about the combinations of mental representation without considering how the syntactic rules bear on the meanings of expressions. This question is somewhat more pressing. It asks how the meanings of pictures and other representations could possibly be combined with those of linguistic representations to form new meanings.
There is a general version of this problem that arises for language-language combinations just as well as it does for other LoT combinations. How do the meanings of ‘is mortal’ and ‘Aristotle’ combine? If this is a problem for this view, it is also a problem for any LoT view (and any language), since all such views posit combinations of meanings of mental representations. There are various strategies one might adopt here, including type theory. I will assume there is some solution to this general version of the problem.

However, the concern about the possibility of semantic combination of pictures and other representations with connectives pushes in the direction of a more serious worry. It is not clear that a truth-functional semantics, familiar from natural language, could be adapted to such combinations. For example, could a connective concatenated with two pictures really be AND? AND is truth-functional. But it is not at all clear that pictures, maps, analogue magnitudes, or vectors have truth conditions in the first place. At best, such representations would be false or meaningless in all but a slim minority of cases. In virtue of even a tiny blemish, a portrait of Macron might fail to be “true” of him. Some have held that it is not possible for pictures to have truth conditions though cf. even in principle (Fodor 1975: 180-1; cf. Goodman 1968, Kulvicki 2006). So perhaps a portrait of Macron could not ever be true of him. Rather, such representations seem better described as accurate or inaccurate, notions that admit of degrees (Burge 2010, Ch. 1; Greenberg 2015, p.2). This would make any connective-like representation that combines with pictures substantially different, in terms of its semantics, from AND.

---

22 A version of this problem is as old as semantics. Gottlob Frege and (especially) Bertrand Russell seem to have been deeply worried about a version of this problem as it applies to propositions, the unity of the proposition. In fact, Hanks (2015) has argued much of Russell’s work in the theory of descriptions was motivated by resolving this problem.
How might one characterize the meanings of cosmopolitan LoT strings to accommodate
this worry? One might simply allow here that many well-formed strings in cosmopolitan LoTs will
be false, but not meaningless. After all, many our mental states (though surely not most of them) are
false or inaccurate. But the proponent of such LoTs might want a semantics that is less
counterintuitive, one that captures the fact that slightly inaccurate representations can still get
important things right about the environment, and can still effectively guide an animal’s behavior.
One might appeal to a different tool for the job: accuracy-functional connectives, which permit
combinations with pictures, maps, and the like possible. Consider AND*, a connective takes the
accuracies of the states it combines and delivers another accuracy. For example, it might deliver
some weighted average of the accuracies of the representations it combines. Or, it might deliver the
lower accuracy of the two. This works for pictures, maps, magnitudes and vectors. Moreover, since
the veridicality of a sentence can be understood as an accuracy of 0.0 or 1.0 out of 1, AND* might
also yield accuracies for combinations of sentences. This is only a cursory description of the
semantics of such connectives. And it bears important dissimilarities to the semantics of most
languages. If a satisfying truth-functional semantics for these kinds of combinations is not
forthcoming, this difficulty might motivate an additional constraint on what ought to count as a
LoT.

---

23 Thanks to Kate Ritchie for help with this discussion.
24 This possibility highlights a syntactic curiosity of AND*: If AND* may be combined with sentences, it can be applied
recursively in at least some cases. If it cannot, then there is an upper bound to the sentences with which it can combine.
If Pic1 and Pic2 are pictures, the string Pic1 AND* Pic2 cannot be combined with AND*, since it is not a picture.
25 Though, interestingly, not at least one previously proposed LoT. See Goodman et al. 2015 for an implementation of
probabilistic concepts in a LoT.
26 NB: There is a substantial literature on pictorial semantics (e.g. Kulvicki 2006, Greenberg & Giardino, eds., 2015;
Greenberg 2015, Wang 2016, Dorit 2018). However, to my knowledge none of these approaches has attempted to give
an account of the semantics of picture-language combinations.
Objection: This view is too liberal in what it characterizes as languages. I claimed in the previous section that a LoT could possess some forms of combinatoriality, but lack others. Very plausibly, a collection of representations that lacks any form of combination of symbols is not a language at all. So maybe possessing combinatoriality of all kinds is a necessary condition on some such collection’s being a language. If that is right, a system that lacked all three types of combinatoriality would not be a language. And so one might question whether a system that has one or two, but not all, types of combinatoriality is a language at all. And if something is not a language at all, it cannot be a LoT.

Reply: This is really an objection to my characterizations of LoTs in Section I. I will focus on the case of languages that lack sub-sentential and sentential combinatoriality, but have suprasentential combinatoriality. A simple example is a language that permits disjunctions of pictures. I choose to focus on this kind case because it will be of interest in discussing empirical results in later chapters. But similar replies to the one I will give for this kind of LoT can be made for other kinds of incompletely combinatorial languages.

One reason to think that LoTs with only suprasentential combinatoriality really are languages is simply that they are not pictures, maps, vectors, or analogue magnitude representations. These are the most familiar kinds of non-linguistic representations. So if these kinds of representations are different from the ones described, then it is much more plausible that they are languages.

27 One might think not: Maybe combinatoriality is not a necessary condition on being a language. In fact, it only matters for my argument that it be sufficient for being a language, and that seems vastly more plausible.
Start with pictures. Plausibly, pictures represent their subjects by means of resemblance (Lopes 1996, Schier 1986). For example, portraits of President Macron resemble Macron. And that is why they are portraits of Macron. If that is right, it rules out the possibility that a combination of pictures with connectives could be a picture. The constituents of the representation would be pictures. But the complex string cannot be. This is because they do not typically resemble what they represent. Consider, for example, a disjunction of two pictures: A picture of Macron giving a speech and a picture of Angela Merkel writing at a desk, on either side of an ‘or’. Though the individual disjuncts might resemble the heads of state (and actions, and so on) that they depict, the complex representation typically will not resemble whatever the disjunction represents. It is difficult to see what, if anything, could really resemble that. So, whatever such a disjunction is, it is not a picture.

Maps, like pictures, represent in a way different from language. Spatial relations between features in a map represent spatial relations between locations in the area represented by the map (Rescorla 2009a). For example, if El Paso is east of Los Angeles, it will be to the right of it on a north-oriented map. But it is sometimes thought that they also have a minimal, language-like syntax, too (Casati & Varzi 1999, Camp 2009). So maps might seem like a more plausible home for our partially combinatorial systems of representation. But the types of sentence that we are considering do not represent as maps, either. Consider complex representations in a system of representation that includes logical connectives. For example:

(3) R1 OR R2

28 Some (Goodman 1968, Kulvicki 2006) deny this. They claim that pictures have a constituent structure, and that picture-constituents represent by convention. On such views, the constituents of pictures are the smallest discriminable bits of pigment in them. But the whole point these authors make here is that pictures are linguistic representations. So if they are right, then claiming that the kind of representations we are considering are pictures entails that they are languages. This serves my point about these systems.
(4) R2 OR R1

where ‘R1’ and ‘R2’ are placeholders for any kind of representation one likes: pictures of heads of state, maps of Los Angeles and El Paso, one-word sentences, traffic signs, and so on. No such disjunction, even one composed of maps, could represent as a map. That is because, as with the constituents of (1) and (2), spatial relations of the disjuncts in (3) and (4) do not affect how any spatial relations are represented by (3) or (4). For example, if R1 and R2 are maps, it just cannot be the case that the places mapped in R1 are both to the west of the places mapped in R2 and also to the east of them. And yet (3) and (4) have the same truth conditions. Similar examples can be constructed for other logical connectives, and for any representations that combine with whole sentences.

This argument is quite general, and versions of it apply equally well to analogue magnitudes and vector spaces. The conclusion we should draw from this is that the kinds of representation that we have been considering are really quite different from paradigm cases of non-linguistic representation. So, if we hesitate to consider the systems that result in these kinds of representations “languages” because those systems seem different from paradigmatic languages, then we ought to hesitate to call them pictures, maps, magnitudes, or vectors, for the same reasons.

It might seem too hard on advocates of non-linguistic representations to say that non-linguistic representations must resemble the things that they are about. This might be true of portraits, which are made (in the first instance, anyway) to be perceived. But surely, the thought goes, nothing in the mind really resembles what it represents. Indeed, who would be looking at them...
(in the first instance, anyway) to determine that they resembled the things they were supposed to represent?

Resemblance actually is a feature of some kinds of non-linguistic mental representation, so there are some formats to which this worry does not apply. Analogue magnitudes and vectors, in particular, change in size or amplitude in a manner correlative with the things and properties they represent. So the kind of argument targeted directly at external pictures and maps should work just as well for analogue magnitudes and vectors. To the degree that a LoT syntax might permit the concatenation of analogue magnitudes or vector spaces with other mental representations, these could be part of the attending LoT.

If we return to picture-like and map-like mental representations, things become a bit more complicated. Surprisingly, though advocates of picture-like representations abound (Kosslyn, Block, Burge, Carey), shockingly little is typically said about what constitutes pictures in the head. A fairly commonly accepted negative definition is due to Jerry Fodor (2007). This definition is meant to be applicable to external pictures, but also to mental representations. Fodor points out that language-like representations, but not others, have what he calls a canonical decomposition: a standard way of splitting them up into parts. For example, consider: ‘The 7-11 is three blocks down on your right’. That sentence has a canonical decomposition, the first step of which splits it into: ‘The 7-11’ and ‘is three blocks down on your right’. A non-canonical decomposition would be ‘The 7-11 is three’ and ‘blocks down on your right’. And a very non-canonical one would be ‘The 7-11 i’ and ‘s three blocks down on your right.’ This applies to sentences, of course, but it also applies to simpler expressions, like noun phrases. For example, ‘tasty bone’ has a canonical decomposition: ‘tasty’ and ‘bone’. By contrast, though pictures, maps, and diagrams have parts, they do not have canonical
decompositions. The case is clearest in pictures: Any part of a picture is a picture of a part of the scene that the larger picture depicts. If I slice out a bit of the middle of that Macron portrait, I will be left with a smaller picture—e.g., of his nose.²⁹

I’ve now argued against the objection that some of the “sentences” in the LoT I have described are not really instances of language by arguing that they are not pictures, maps, magnitudes, or vectors. But might they be a sui generis kind of representation?

If the kind of representation I have described really is a sui generis kind, it is very similar to linguistic representations, because it is partly combinatorial. Call such kinds L*OTs. The arguments in this section, and elsewhere in this dissertation, would then be aimed at this slightly modulated kind. I see no problem with concluding that the arguments of this paper are really about L*OTs. These bear striking and important similarities to LoTs. (And, so far as I can tell, the LoTH is as compatible with L*OTs as it is with LoTs).

This points to a surprising possibility: that a system of representation’s being combinatorial is not sufficient for its being linguistic. The dialectic so far has been: Collections of representations are linguistic if they are combinatorial; pictures, maps, etc. can be combined; therefore some languages include pictures, maps, etc. Rebuttal: Those just can’t be languages, because they are not combinatorial enough. But if that is right, then the family of combinatorial representational systems includes languages, but is not exhausted by them. And if that is right, then distinguishing between

²⁹ Might Merkel’s painting can canonically decompose? For example, might it include a painting with her pen as a constituent? If it turns out that there really is not a good way to distinguish linguistic from non-linguistic representations on these grounds, and there is no plausible alternative, then this would seem to favor the proponent of cosmopolitan LoTs: “Non-linguistic” representations ought, by assumption, to be combinable, since it is not really in dispute whether linguistic ones are.
varieties of combinatoriality in the way we have done so far still amounts to progress in our understanding of LoTs, and of linguistic representation more broadly.

Of course, whether one wants to go in for such a claim depends on how closely one holds to a particular picture of what makes language special. If one prefers the idea that combinatoriality is unique to language among representational formats, then one ought to go in for the kind of view I have been describing in this section. (There is some reason for doubt: syntax tree representations are combinatorial: branchings serve to combine nodes. Plausibly, so are mental representations of at least some dominance hierarchies (Cheney & Seyfarth 2008; Camp 2009) and preference orderings). But one might prefer to take on the idea that combinatoriality is merely a necessary feature of languages. Genuine languages require full-throated combinatoriality across (at least) the three varieties described earlier. Thus, The Language of Thought Hypothesis was misleadingly titled, since some languages of thought are not languages. There might also be other features of language, beyond mere combination, that one sees as necessary for language: the syntactic capacity for recursion, or a syntactic embargo on picture-like or map-like representations. A radical position might simply drop the idea that language is that special: Natural language, which by most accounts is recursive and combinatorial nearly across the board, is merely a point very near the end of a scale, along which fall many other means of linguistic representation.
Chapter Three: A Lack of Evidence for Non-Human LoTs

In this chapter, I survey several arguments for the LoTH, and consequently for the existence of LoTs. Some are familiar from debates about the LoTH, and some are, to my knowledge, novel. I argue that they all fail when applied to non-human animals. So, even if it should turn out that human cognition includes a LoT, this would not guarantee that any other species does.

Languages of Thought and “Non-Linguistic” Animals

Proponents of the LoTH, including Jerry Fodor (1975, 2008) and Zenon Pylyshyn (2009; Fodor & Pylyshyn 1988), take it to be a boon to the hypothesis that it can explain not just a great deal of adult human cognition, but also a great deal of preverbal human and non-human cognition. Even without considering empirical evidence for their mentality, it is overwhelmingly plausible that some non-human animals have intentional states. And the LoTH can explain how this is possible, despite the fact that non-human animals lack natural language.

The LoTH has been immensely influential, if philosophically controversial, as account of (certain aspects of) human cognition. Yet comparatively little has been done to motivate the claim that a language of thought exists in other animals. The cognitive ethologists Dorothy Seyfarth’s and Robert Cheney’s book, Baboon Metaphysics, is a noteworthy exception (Seyfarth & Cheney 2008). In it, they argue that since our best evidence suggests baboons’ representations of social dominance hierarchies include recursive elements, we should conclude such representations are encoded in a language of thought. They describe many cases of troop-wide changes in behavior in the wake of
interactions that impact dominance hierarchies. A typical example is lower-ranking males attacking higher-ranking males. If the attack is successful, all the members of the group who witness the interaction change their behavior toward other baboons in the troop accordingly. They seem to update their dominance relation representations, not just for those directly involved, but also all those indirectly affected by, the ensuing dominance shift. In addition, when the conflict occurs between members of different matrilines, dominance relations for entire families of baboons are updated. For Cheney and Seyfarth, this kind of phenomenon is best explained by recursive embedding of dominance relation representations. If A, B, C, and D name individuals or matrilines, such a representation might resemble the following: A DOMINATES B, WHO DOMINATES C, WHO DOMINATES D. It is easy to see how the capacity that allows this could allow for embedding of many more individuals or matrilines, suggesting iterability of the DOMINATES concept. But this kind of iterability just is combinatorial. And so, this hierarchy representation updating is supposed to be evidence for a language-like structure to the baboons’ thoughts, at least about dominance hierarchies. This is just the kind of evidence we need to answer the question of whether baboons can think thoughts only possible with a LoT.

In a reply to this book, Liz Camp (2009) contends that cognitive maps account for the baboons’ behavior as well as a LoT might. Cognitive maps, which do not have a linguistic syntax, nonetheless have the representational power needed to explain the baboons’ behavior. Surely if such maps are part of the representational repertoire of the baboon mind, they could explain the cognitive achievements described by Seyfarth and Cheney. One could doubt whether they do occur. But given their extensive role in psychological explanation, especially in explaining navigation, a more conservative approach should be to accept such maps as legitimate psychological kinds and so potentially as a means of representing hierarchy relations in baboons. To overcome this important
objection, one must supply either independent evidence for a LoT in baboons, or independent
evidence that it is unlikely they represent dominance relations by means of maps. The important
thing to recognize, for now, is that as it stands, Cheney and Seyfarth’s argument is merely suggestive of
a language of thought in nonhuman primates.

In the next section, I’ll look at some different strategies for detecting non-human languages
of thought. Before moving on, though, I’ll consider one argument against languages of thought in
non-linguistic creatures. I assume here that ‘non-linguistic’ and ‘non-human’ are pretty near
coeextensive: Though of course there are many non-linguistic humans, there are perhaps no other
genuinely linguistic species. And so, an argument against non-linguistic animals’ LoTs is an argument
against non-human LoTs. Even if one disagrees with this assumption it is usual to think the class of
linguistic creatures is very small. And so, the argument at issue, due to Jose Luis Bermudez (2003),
will be directed to a very similar class of animals to the one I am considering. The argument is
supposed to show that any animal with a LoT ought to have natural language. It goes as follows:
Natural language confers powerful advantages for survival to those who have it. The language of
thought facilitates natural language acquisition. So it is unlikely that any creature that possesses a
language of thought should lack natural language. Yet only humans have natural language. So, it is
unlikely that other animals than humans possess a language of thought—even if human beings do,
in fact, possess one.

This argument conveniently glosses over whether some physical capacities might be required
to produce language. Some animals might simply lack expressive systems, like vocal cords or
dexterous hands, with sufficient degrees of freedom for expression. Additionally, some might lack

30 In fact, for some such as Fodor (1975), it is thought to be necessary for natural language acquisition
the attentional or memory resources to process strings of natural language, which occur across time. These are just a few examples, but they could be multiplied, and they point to a facility in Bermudez’s assumption that the movement from possession of a language of thought to possession of a natural language should be relatively easy.

The idea that a LoT should bring along with it a capacity for natural language seems most natural if one thinks that the primary function of language is communication. But it is not obvious that this is the case, as many have argued (Devitt, Fodor, Chomsky). Language also confers strong advantages to one’s thought. For example, it is plausible that a LoT facilitates deductive inference and recursive thoughts, and that it expands one’s potential conceptual repertoire. Unlike communication, this is adaptive for solitary animals as much as it is for social ones.

Finally, this argument, even if it goes through, does not, all on its own, show that non-human animals lack LoTs. That conclusion is only supported if there are no reasonable independent grounds for supposing that some animal possesses a LoT. So, the argument merely places a burden on the proponent of non-human LoTs to supply positive evidence in favor of such languages. As I suggested earlier, there is independent evidence for this, which we will return to in later chapters.

Some Obvious Ways to Tell if an Animal Has a Language of Thought, and Why They Don’t Work
The obvious strategies of arguing for LoTs in non-human animals will not work. In this section, I detail the reasons why. A first pass strategy is to take traditional arguments for a LoT in humans and see how they fare in the non-human animal case. The strategy meets varying degrees of failure, depending on the particular arguments one chooses. Several of Fodor’s arguments in favor of the LoTH are difficult to apply to non-human animals. Some are best motivated by traits that do not exist in non-human animals. Others rely on traits that are difficult to recognize in animals.

Consider Fodor’s argument from systematicity: Our ability to think that Jim loves James brings along with it an ability to think that James loves Jim. Surely this is true of us. But the most obvious way in which we can figure out that this is true of us is not available in the non-human animal case. One reason to think it is true is because we can introspect our own conscious thoughts. And every relation we can think of is one we can understand inverting. Another reason comes from observation of language: Anyone who can say or understand ‘Jim loves James’ can say or understand ‘James loves Jim’. So I can be confident that your thoughts observe systematicity by listening to your words, and taking on the modest assumption that, by and large, what you say expresses what you think. These are not the only pieces of evidence (see below) one might marshal, and certainly not the most convincing, but they are the most salient ones in the human case.

The problem with these first methods is that we are not in a position to introspect or even really imagine the thoughts of non-human animals, and non-human animals lack natural language. Regarding introspection, we are actually in a much worse epistemic position with regard to the mental states of non-human animals than we are with regard to those of other people. For example, an inference from the apparent structure of my thoughts to the structure of my friend’s thoughts is stronger, all else equal, than an inference from the apparent structure of my thoughts to the
structure of my dog’s thoughts. And of course, regarding linguistic evidence, we cannot infer the presence of systematicity from it if there is none! One can construct similar puzzles for Fodor’s arguments from productivity and natural language acquisition. How do we know that thought is productive? Again, in the first instance by considering our own thoughts, and by observations of natural language. But these routes to productivity won’t work for my dog. And, of course, natural language acquisition is a non-starter for animals that never do acquire language.

But aren’t there other, better ways of detecting systematicity and productivity? The argument below from Fodor and Zenon Pylyshyn, which anticipates the concern of this section, provides an argument in the systematicity case that relies on the systematicity of perception (a parallel argument can be constructed for productivity):

It is not, however, plausible that only the minds of verbal organisms are systematic. Think what it would mean for this to be the case. It would have to be quite usual to find, for example, animals capable of representing the state of affairs aRb, but incapable of representing the state of affairs bRa. Such animals would be, as it were, aRb sighted but bRa blind since, presumably, the representational capacities of its mind affect not just what an organism can think, but also what it can perceive. In consequence, such animals would be able to learn to respond selectively to aRb situations but quite unable to learn to respond selectively to bRa situations. (So that, though you could teach the creature to choose the picture with the square larger than the triangle, you couldn’t for the life of you teach it to choose the picture with the triangle larger than the square.) (Fodor & Pylyshyn 1988)
I think this argument gets us part of the way, but there are some ways in which it falls short of really reassuring the proponent of non-human languages of thought. First, one needn’t have concepts that map on to each representational type one can token perceptually. Vision represents numerous features of the environment for which only those interested in vision science have concepts. When I was fourteen I lacked the concepts T-JUNCTION and CONTRAST EDGE but I am extremely confident my visual system tokened the corresponding visual representations anyway. So, it is not crazy to imagine that the relations represented in perception will similarly come apart from those represented post-perceptually. Second, relations of size and color, which Fodor and Pylyshyn cite, are archetypes of perceptual features. And so the explanation of such cases could be cast entirely in terms of perceptual learning using these perceptual features. In such cases, it is an open question whether the animals have mental states that are not perceptual in nature, and whether those states are best explained in terms of a LoT. So, even if this argument goes through, it only pins the putative LoT to perception (or, perhaps, in motor commands: those, too, partly explain the animals’ behavior, and might be encoded in a LoT). And that matters for us here, since we know that many animals have mental states that are not perceptual (or motoric) in nature. Second, one can only imagine that the number of animals that have explicitly been tested on learning tasks like the one described is small. I do not know of any, and Fodor and Pylyshyn do not cite any. So the claim is, at best, speculative for most animals. And if you are disinclined to think of non-human animal mental states as systematic, because for example you do not think they possess a LoT, you will be likely find the speculation that non-human perception respects systematicity less plausible.

Another problem with this style of argument is most visible in the parenthetical at the close of the citation above. Associative learning is an initially very appealing means of detecting exceptions to systematicity in nonhuman animals. Asymmetries between one’s ability to represent aRb and bRa
should manifest in learning, where they will generate asymmetries in behavior. But there is nothing unusual about the fact that animals can learn some relations and not others. Rats, for example, are prepared to associate gustatory stimuli with nausea, but contraprepared to associate them with audiovisual stimuli. This should not be taken as evidence for a lack of systematicity. Rather, it should be taken to show that evolutionarily prepared responses generate exceptions in learnability. It might well be that in some cases learning aRb is, for similar reasons, more difficult than learning bRa—without providing us any reason to suppose systematicity has broken down.

Other arguments for a LoT apply more straightforwardly to non-human animals. Unfortunately, the road from the conclusions of these arguments to animals is still rocky. First, there is the argument from perceptual representation. Fodor begins with the claim that perception is computational, a central tenet of the study of human perception. I find this completely unproblematic (cf. Orlandi 2014). Similarly unproblematic is the idea that this brings along with it the claim that perception is representational, since as Fodor claims, there is “no computation without representation” (Fodor 1981, p. 180). A mass of evidence shows that many animals have genuinely representational perceptual systems. To take some examples nearly at random: Pigeons are sensitive to modal completion cues in some contexts, but not all, while chimpanzees seem to be sensitive to them in all of the situations that we are. Several fish species appear to be sensitive to versions of the Müller-Lyer and Kanisza figures. There is even an emerging psychophysics for bat sonar, replete with interpretations which appeal to representations of spatial properties (e.g., Akre et al 2011; Warnecke & Simmons 2016). These examples could be multiplied. I take them together to make the premise that some non-human animals perceptually represent the world to be beyond

31 Mandelbaum 2015 offers a longer discussion of these issues for associationism (I credit Eric, and that entry, for this argument).
reasonable dispute. So if this argument is good, it really does show some animals have a LoT, and in fact, a great many do.

But from there, one must specify the kind of the representations that are being computed over. Here, things begin to be more tendentious. An underlying premise here seems to be that the only possible format for perceptual representations is a linguistic one. But it is plausible that at least some perceptual representation is picture-like, rather than language-like, as Fodor himself and others later observed (Fodor 2007, 2008; Quilty-Dunn 2016; Block unpublished manuscript). And if there are non-linguistic perceptual representations, we would need to show, at a minimum, some positive evidence that those representations could be combined syntactically. There is a possible hybrid view, according to which there is a co-habitation in perception of picture-like and language-like representations (Quilty-Dunn 2016). This avoids that problem, since of course the language-like representations will have combinable parts. (However it does raise a new question of whether those interface with one another, or how they might both play a role in the same perceptual system if they do not). The important thing to recognize, here, is that the computational nature of perception does not give us for free that perception occurs in a language of thought. One needs independent arguments for that.

At least some aspects of perceptual processing likely include combinations of representations. But is worth stressing that even if that is the case, this still tells us nothing about thought. There are (and must be) independent arguments for LoT's in cognition. That is because for all we know, thought and other aspects of cognition might work in importantly different ways from perception. My dog’s perceptual systems might be modular, for example, and his learning and decision making might be entirely determined by associative links between, or Bayesian updating
processes (Rescorla 2009c) over, the outputs of those perceptual systems. The language of thought hypothesis is misnamed in that it is supposed to characterize not only thought, but also perception and most things mental. But it is supposed to characterize thought. If it does not do that, this is a shortcoming of the view.

Another more promising route is the argument from concept learning. Fodor’s view (2008) is that models of concept learning require hypothesis testing and confirmation. And in order to do this, one must have some means of representing the concept in one’s hypothesis. Any concept one can learn must be a concept one can build—it must be something expressible by one’s existing conceptual repertoire. But then, for any concept c one could acquire, one already has the means of conceptualizing c. This brings along with it an indirect argument for LoTs: In virtue of what could one conceptualize c, but a language with some compositional structure? One learns c by concatenating other concepts; the tool for concatenation is language.

One might worry about how we should characterize the concepts of non-linguistic animals (Stich 1983) in the first place, or what the right tests of category learning should be for them. If so, then concept learning in non-linguistic animals might be too much of a mess to even take up as a means of determining that animals have a LoT. I will for take for granted that we can determine the intensions of animal concepts, as well as when an animal has acquired a given concept. Some promising strategies for this include the extension of habituation paradigms, which have been used on very young children, to non-human animals (Laurie Santos’ lab at Yale has started to apply this method to rhesus macaques, but the work is not yet published.) And anyway, there is ample reason to believe we are ignorant of many of our own concepts, and those of others. And so if one is
genuinely worried about this problem for non-human animals, then one’s problem is far more general than concerns us here.

Another potential worry is that the LoTH does not offer the only serious account of concept acquisition, or of concepts; Fodor’s argument does not start from widely accepted claims. Many who like the idea that prototypes are concepts will still have a strategy for accommodating the compositionality of concepts (for a useful review critical of these strategies, see Gleitman et al. 2012); anyone who prefers the hypothesis that Quinean bootstrapping explains the acquisition of some concepts (Carey 2011) is going to take issue with hypothesis testing and confirmation. Still, one need only accept Fodor’s picture of concepts and concept learning, and one gets for free that any animal that can learn a concept has a LoT.

But there is a more general shortcoming of this argument: It can only provide direct evidence for an incompletely combinatorial LoT. If Fodor is right about concepts and how they are acquired, and this picture applies to at least some animals; then those animals have the ability to compose concepts, such as TASTY and BONE, to make others, such as TASTY BONE; and very plausibly, they also have what they need to make sentences such as the thought THAT IS A TASTY BONE. But there is no evidence from this picture of concepts that any non-human animal will have supra-sentential combinatoriality: They may not be able to form mental representations that include combinations of representations with entire sentences. This includes combinations of sentences with logical connectives, such as: THAT IS A TASTY BONE OR THIS COUCH IS A GOOD PLACE FOR A NAP. And this would seem to matter, since one reason to care about LoTs is that they might underpin not just our concepts and our thoughts, but also many of the transitions between those thoughts. An
important aspect of combination—one that is partly responsible for deduction—is left unaccounted for, even if this argument is good.

So far in this section, I've canvassed several traditional arguments for the LoTH. The goal has been to show that they do not work very well when applied to non-human animals. I will now look at an argument that has not been often adopted by proponents of the LoTH, but that nonetheless might at first seem appealing. It relies on inferences from what we know about human cognition. It might seem that, lacking evidence to the contrary, we ought to assume that phylogenetically close species are likely to share traits. So, animals that are phylogenetically close to us are likely to share traits with us. And if those traits include a LoT, we have some reason for supposing at least some other animals have one. This is especially appealing for close relatives of ours, such as chimpanzees and bonobos. Such a strategy is, in effect, just a rejection of fears of anthropomorphism in comparative cognition. A toy example might go as follows: We don’t have evidence against a language of thought for bonobos. And we do have evidence for one in humans. So it is prima facie plausible that bonobos possess a LoT.

This is only the sketch of a view but it is already easy to see there are serious problems with it. For one thing, evolution does not seem to be nearly as accreditive over generations, nor as gradual across clades, as we’d like for this kind of inference to be at all reliable. It is possible that something as consequential as a LoT could have emerged quickly in evolutionary time. This is to say nothing of the fact that we killed off the species most closely related to us, and so the gap between us and our closest surviving neighbors happens to be rather wide. Another issue is that it is hard to specify an acceptable notion of cladistic closeness that would be helpful beyond such obvious cases. It will never be vague whether an animal has a language of thought; having a language of thought just
is not the kind of property that lends itself to Sorites paradoxes. (That said, as mentioned earlier, it is possible that a species could possess some of the features, but not all, of a LoT, in which case there might be unclear cases). And finally, it will not do to stipulate a notion of ‘closeness’ here: Whether an animal has a language of thought is an empirical question and just cannot depend on the working definition we choose for ‘close’. A well-developed comparative psychology might produce patterns in historically shared traits that help us put together an informed definition of ‘closeness’. But so far as I know, no definition of this kind is forthcoming. And anyway, an important aim of comparative psychology is to recognize aspects of thought that are shared by humans and non-human animals. So it would be curious if a well-developed comparative psychology needed such roundabout means of identifying the presence of the language of thought in non-human animals.32

So far, I’ve canvassed several ways one might try to identify LoTs by way of a language of thought in non-human animals. Many of the traditional arguments for the LoTH in humans are difficult to apply to non-human animals. It won’t do to single out species that are close cousins of humans. So, though it seems uncontroversial that many animals perceive, and that some think, decide, and act, it is not obvious they do so by means of a LoT.

Looking forward, it is worth wondering what a reasonable argument in favor of non-human LoTs might look like. There is a final, indirect strategy that can also be attributed to Fodor, but that is much less often cited than the arguments considered in this chapter, that is worth considering. This is the idea that cognition is an interaction effect “par excellence” (1983, p1), a fact that only

---

32 Interestingly, intuition pulls us in the opposite direction in the case of innate concepts. Likely, we possess innate concepts for adaptively crucial categories, such as FOOD, CONSPECIFIC, or MATE. And there is little reason not to suppose that some other animals possess these concepts—or similar ones—too. Depending on the kind of theory of concepts one endorses, and also on the way one chooses to individuate concepts in non-human animals, this might be an avenue toward attributing LoTs by means of the method just canvassed.
LoTεs can really explain. My beliefs and other mental states bear intricate relationships to one another, and to my behavior. The only way we might hope to explain my behavior is in virtue those relationships. And the way to explain those relationships is in virtue of semantic relations. The LoTH allows for inferential relationships, and so for semantic explanation of these interactions between mental representations, because it posits a language for those interactions. Thus it can explain my behavior.

The strategy for identifying animal LoTεs with this method is to seek out animals with behavioral complexity that could only be explained by such intricate semantic relations. In the human case, it is abductive, starting from the patent complexity of human behavior, and working back to its underlying cause. So, as a diagnostic tool, it is quite difficult to apply to any particular species. It requires extensive, prolonged investigation of a range of behaviors. A less painful, more productive strategy will be to appeal to a range of species, and to show that the same tool has impressive predictive power for a limited number of observations of each. Moreover, only a LoT with a lexicon that includes sentential connectives, quantifiers, or some other operator that can plausibly make possible inferential relations, will display the “interaction effects” in behavior that Fodor describes. So the strategy implies that one identify such operators. This strategy has, to my knowledge, never been applied. It is the aim of Chapters Four and Five.
Chapter Four: Minds Without Spines: Mental Representation in the Fruit Fly, and Arthropod Languages of Thought

So far, I have described some different ways that a LoT might be, and shown that there are few if any good reasons to think any non-human animals have them. This chapter, and the next, will provide some positive evidence that at least some non-human animals are capable of combinatorial mental representation. Here, I look at minimal cases. First, I consider a case of an animal that mentally represents but—for all we know—cannot combine its representations syntactically: the fruit fly. To do this, I employ an empirical strategy for identifying the presence of mental representation in other animals that has been underappreciated in the philosophical literature: tests of multi-stable perception. Then, I discuss seemingly simple creatures—bees and wasps—that behave in a way that suggests they form complex mental representations of nectar sources. I conclude with some considerations about the syntax of their LoTs.

Mental Representation in the Fruit Fly

We can start with the case of mental representation in the absence of combination. The roadmap here will be to (a) demonstrate that a particular kind of test is diagnostic of mental representation, and (b) point to evidence that an apparently simple arthropod meets it. We have no reason to suppose this animal combines its representations syntactically. Thus, it provides tentative evidence of the existence of minds without LoTs: Mentality appears to be more widespread in animals than combinatorial minds are.
First, I propose the diagnostic test for mental representation. To pass this test is to meet a criterion for representing at all—for being the kind of thing that has mental representations. For our purposes here, this is important, since mentally representing is necessary to possessing a LoT. But there is also ample reason to care about finding such tests, independently of this consideration. First, it’s scientifically useful to have a diagnostic tool of this kind, and as many of them as possible. Such tools can help us recognize creatures that are capable of more than merely registering changes in stimulation of their bodies. It can help us recognize, instead, the animals that perceive distal objects as such. And second, it can offer us part—though not all—of an account of how to draw a philosophically storied boundary: that between sensation and perception. An animal with sensory transducers, such as retinas or olfactory nerves, but no mental representations, lacks mind in just the way that a weather vane or a thermostat does. The discovery that an animal has or lacks mind is important, and so how this line is drawn is important, too.

I do not take the criterion I am going to give to be the only indicator that an animal has mental representations. In fact, it’s plausible that some creatures might fail to reach the criterion, though they have mental representations of some kind. But I do take it that a creature’s (or machine’s) meeting the criterion is sufficient for mental representation.

I use ‘mental representation’ here in a way compatible with most theories of mental representation. I focus on a familiar and plausible condition on representing, and one that any theory of representation should accept: With mental representation comes the ability to get things wrong. In accurately seeing a plum as green, it must be the case that I could have been in just the same mental state, though the plum were not green. In falsely judging that plum trees cannot grow in central California, it must be the case that I could have been in just the same mental state, though
central California really did turn out to be, for some reason or another, unsuitable for plum trees. And of course, that the plum not be green and that California be plum-friendly should be part of what makes it the case that those representations are true or false. This is just part of what it is to represent at all. So long as one is comfortable with describing minds at least partly in terms of mental representation, I take it that this minimal use of ‘mental representation’ needn’t be committed to any particular theory of representation.

The criterion I propose is sensitivity to multi-stable stimuli. Multi-stable stimuli have been recognized by philosophers and psychologists since at least the eighteenth century (Dutour 1760; attributed by Schwartz et al. 2012) and have been a source of intrigue for philosophers interested in perception (most notably, Ludwig Wittgenstein). They are a staple of vision science research today. And I'll argue that susceptibility to them is evidence that a creature is capable of getting things wrong, and so of mentally representing. As such, it can be a diagnostic tool for the presence of mental representation.

Multi-stable images (or videos) are single stimuli that tend to yield multiple visual interpretations of those stimuli in the same perceiver (This assumes, of course, no eye-movements, squinting, or other perturbations of the proximal stimulus by the perceiver. Anyone can alternate their perceptual state by changing their perceptual inputs). In human vision, examples include the Necker cube, the duck-rabbit, bi-stable motion, and binocular rivalry, among others. Typically, one can choose to switch between visual interpretations if one wishes: One can attend to the rabbit's nose to see the rabbit. But multi-stable images also lead one to alternate between interpretations spontaneously. For example, most people inevitably undergo a gestalt switch to rabbit after staring at duck long enough. Multi-stable images are plausible indicators of mental representation partly
because this stochastic response is so hard to account for without appealing to mental representation. The reason is straightforward. No change in the stimulus has occurred, and yet the creature’s behavior toward the stimulus changes. Specifically, the creature behaves as if it is now in the presence of a different stimulus than some time before. But of course, it isn’t! (I will say more about this later).

It should be clear now why I do not bill this as a necessary condition on representation. Some animals might perceptually represent, though for whatever reason, their visual systems never do so in a multi-stable way. This might be the case in some very rudimentary forms of perceptual representation. It is an empirical question whether any extant perceiver does function this way (I doubt that any do), but so far as we know, it is possible one could. I say “so far as we know” because there is no obvious reason to think such a creature’s behavior would not support many of the counterfactual claims we would want it to support if it were a creature who takes its environment to be a certain way.33

Care must be taken here, of course, to distinguish between creatures that visually represent their environments from those that are behaviorally responsive to changes in their environments but don’t represent them. A creature’s sensory transducers might be sensitive to certain distal changes,

33 There is actually some room for doubt here. Take any visual representation type \( r \). Suppose, as is likely, some set of computations, \( z \), typically eventuates in \( r \) (there may be several computational routes). In that case we can take the set of those sets. It seems plausible that whatever \( r \) may be, one could generate some visual input, \( i \), under which the visual input to \( s \) strongly supports tokening \( r \), but also strongly supports tokening some other represention, \( r' \). But what if the design of the visual system in question is absolutely asinine, and can output only one kind of visual representation, or very few? In much the same way, one could generate some \( i' \) such that \( i' \) only weakly supports tokening \( r \). Perhaps it is completely uncertain, as far as the visual system is concerned, whether \( r \) obtains. Either route seems highly likely to lead to multi-stability. Whether the computations underlying the representation of any features in such simple creatures is susceptible to such cases is an open question. But all this does give us some reason to suppose that any perceiver is in principle a multi-stable perceiver. So while I will only stake a claim to sufficiency here, I am tempted by the necessity claim as well.
and those changes might even lead to behavioral responses to those changes, even though that sensitivity and response are not underwritten by anything like a mental representation. This is plausibly what occurs in the case of the phototropic paramecium (for a discussion, Fodor 1987) or the sunflower, both of which respond to changes in the distribution of natural light in their environments. The paramecium “swims” away from natural light; the sunflower orients toward it. Recent discussion of this distinction include Burge (2010) and Rey and Knoll (2017). The former argues that perceptual constancies—the ability to track objects and their features across changes in retinal stimulation—demarcate the perceptual; the latter argue that genuine representation requires being able to recover from errors in one’s perceptually-guided actions. Since the criterion I am giving is a sufficient one, it must be that any creature that meets it is capable of more than mere sensation. But this criterion is compatible with the claim that the true lower bound between sensation and perception resides elsewhere than in multi-stable perception. I refrain here from committing to an account of just where that boundary lies. For all I will say, it could turn out that sensitivity to multi-stable stimuli just is the boundary. Even if it is not part of what constitutes the boundary, there is some reason to believe it is tightly connected with things that are.

As an example of a case in which multi-stability is tightly connected with a characterization of the lower bound of perception, suppose Burge is right that perceptual constancies mark the boundary between sensation and perception. Constancy involves getting away from the stimulation of transducers (the pattern of stimulation on the retina, for example), and holding fixed certain features of objects, independently of changes in sensory stimulation. In one kind of color constancy, for example, an object is represented as being of the same color before and after passing into a shaded area. This happens despite the fact that after passing into shade the object now reflects a different range of visible light. It seems likely we will be able to generate multi-stable images in
nearly any system that is capable of producing perceptual constancies. To take the same example of human color constancy, one might exploit the parameters of color perception to generate stimuli which are ambiguous (as far as concerns the visual system) between shading and non-shading contexts, leading to stimuli that are bi-stable with regard to color. Arguably, something similar to this is going on with “the dress” (though it is not a case of bi-stability, strictly speaking, since most people settle on one interpretation, and struggle to see the other. The differences in color perception show up between, rather than within, individuals).

The kind of switching between representational states typical of the perception of multi-stable stimuli arises primarily in perception. It does not occur in paradigmatically cognitive aspects of representation, like belief or judgment. So, is this really a mark of representation tout court, or is it simply a mark of perceptual representation? If multi-stable switching is merely a mark of perceptual representation, perhaps creatures could mentally represent without perceiving. In that case, this criterion would turn out to be diagnostically imperfect because of the possibility of type two errors (false negatives).

A presupposition of this question, of course, is that there is some non-trivial distinction to be had between perceptual and cognitive representation. Some will reject that presupposition. For example, if the LoTH is true of both perception and cognition, as Pylyshyn and, to a lesser extent, Fodor, have argued, then there is no difference to be had between perceptual representation and other kinds of mental representation—just a distinction between perception and cognition. (NB: One need not defend this claim to defend the LoTH, since the hypothesis might be true of some parts of the mind, such as thinking and reasoning, but not others, such as perceiving). But then
again, some will happily accept that presupposition (Burge 2010, Carey 2011, Block unpublished manuscript).

In any case, it is doubtful there are any such creatures out there—at least on earth—since perceptual representation is ancient. Perceptual representation in animals is likely to be nearly as phylogenetically old as self-propelled movement. On the assumption that all creatures that mentally represent are at minimum capable of perceptual representation, perception is a reasonable place to look for representational capacities. This is most useful in the relatively simple creatures about whom doubt with regard to representation most often arises. I acknowledge the conceptual possibility that a system could exist that mentally represents but lacks percepts. One might, for example, think this is what computers do. And Descartes seemed to believe this is what human minds actually do, once no longer enmeshed with a living body (Descartes 1649/2009). So at least as regards non-human animals, the assumption that creatures that lack perceptual representation lack representation tout court is quite reasonable.

In the rest of this section, I will discuss a case from a model insect species in biology: *Drosophila*, the fruit fly. I’ll argue that since *Drosophila* bi-stably respond to at least one kind of stimulus, they perceptually represent. Thus, the fruit fly has rudimentary mind. This conclusion shows that even seemingly simple creatures may have minds. I also provides a general template for the kinds of experiment that might use the criterion I’ve just laid out to detect mental representation—one that has been underexplored in the philosophical literature on perception. Finally, and most importantly for our purposes, the result is instructive for theorizing about LoTs. Since, for all we can see, fruit flies lack the ability to combine mental representations, they may offer a case of creatures that possess minds but lack LoTs.
In a brilliantly titled paper, “Fruit flies are multi-stable geniuses”, Pack and Theobald (2018) describe a study that relies on the tight link between perception and action in fruit flies to test for sensitivity to a novel multi-stable stimulus. They find that fruit flies act in a way that suggests they multiply interpret scenes of noisy motion. So according to the criterion I’ve laid out, they have representations of motion. I will describe that experiment and say why it—and experiments like it—could be used as a marker of representation. The study on *Drosophila* makes a useful case study for future applications of the method.

Before getting to the design of the experiment, here is a usefully simplified description of fruit fly navigation. A fruit fly’s flight is mostly guided by its vision. *Drosophila* has eyes fixed in its head. So optic flow—the uniform movement of light across the fly’s many eyes—provides the fly with a fairly reliable indicator of motion: Rightward optic flow is a reliable indicator the fly’s whole body is moving left; leftward motion, that it’s moving right. As it happens, when optic flow reaches a threshold of coherent motion in *either* direction, the fly “corrects course” by flapping its wings in a way that re-orient itself in the opposing direction.

If this were *all* the information we had to go on about the fruit fly, a simple, non-representational system might seem to explain how fruit flies maintain course in a changing environment. The proposed system would pair vectors of optic flow to motor system commands. Plausibly, once optic flow reaches a threshold of uniformity and speed across the fly’s eyes, the motor system would respond with “counterbalancing” wing flapping. As a result, the fly’s body would correct course. Such a system might allow the fly to maintain forward motion in a changing environment without ever visually representing that environment. (This is not, in fact, what is going
on. But given only what I have said so far, it could be). Of course, the eyes’ registering of information here might be considered representational in a minimal sense. However, this minimal sense of representation is just that banal sense in which a sunflower or a paramecium “represent” the presence of sunlight. The present question is whether there is some reason to think *Drosophila* represents in a more robust sense than the paramecium or sunflower does.

The study’s authors wanted to see if a deflationary explanation of the kind described in the previous paragraph could be separated out from a representational explanation. They relied on the tight link of stimulation of the fruit fly’s eye to movements of its wings. Here is how the experiment worked. The experimenters fixed flies’ bodies in place with sticks, so that flaps of either wing no longer generated movements of the body, nor accordingly, with movement across the flies’ eyes. They then placed the flies in front of a screen which can simulate optic flow. In other words, the screen could present images that mimic the visual scene, from the fly’s point of view, of leftward or rightward motion. Using this screen, they presented the flies with two overlapping images of wide field motion (motion over a large area of the visual field), with the overall motion in each image going in opposite directions. The resulting stimulus is a bunch of moving dots. Some of the dots move rightward, while others move leftward. This generates a split of leftward and rightward optic flow on the fly’s eyes. If the fly’s movement relies on the kind of non-representational system described earlier, one that links optic flow to wing movements, this should generate a particular kind of motor response. The fly’s motor system should be getting continuous, similarly powerful information in favor of *flapping left* and in favor of *flapping right*. The fly should flap both ways, or (perhaps) not at all. But if, on the other hand, the fly’s vision is using the optic flow to build a more global interpretation of the motion in its visual field, it is reasonable to predict that the fly would respond in a multi-stable way. The competing forms of optic flow are rarely encountered in the wild.
and so plausibly the result of sensory noise. So the visual system might settle on one interpretation of wide field motion at a time. Since there is considerable information supporting both the interpretation of leftward motion and that of rightward motion, the dominant, action-guiding interpretation could be expected to switch, at random, from time to time. (This is what happens for us: after looking at duck at length, one sees rabbit suddenly). Thus, one could reasonably predict the fly would alternate between flapping as if it needed to move left and flapping as if it needed to move right. And this is what was observed: Alternating flapping, sometimes rightward, sometimes leftward. Importantly, the switches happened at random, suggesting a change had occurred in the fly’s vision.

![Diagram of fruit flies with ambiguous motion stimulus](image)

**Figure 1:** Fruit flies were presented with an ambiguous motion stimulus. Reproduced from Toepfer et al. (2017)

---

34 There was, however, a third kind of response, as well: a mix of leftward and rightward flapping. This third behavior response turned out to vary in the percentage of rightward and leftward flapping, for both wings, with the mean motion direction of the stimulus. It is possible then that the response was not an instance of the fly reverting to a simple input-output routine, but of a third representational interpretation. The flies’ perception of the stimulus may have been tri-stable.
It is difficult to see how to spell out such a case for this alternating wing flapping without appealing to perceptual representation. In the cases in which the fly is flapping almost exclusively to one side, it seems to be getting things wrong about the way the world in front of it actually is. It is flapping *as if* it is only receiving visual input typical of leftward (or rightward) motion at any one time. But it is receiving both at all times. And, though no relevant change occurs in its visible environment, it still alternates its motion responses. This behavior is aberrant and calls out for explanation. Moreover, it is straightforwardly predicted by the supposition that fruit flies visually represent motion.

One might think that because of the low ecological validity here, there’s little reason to trust this result to be indicative of the flies’ cognitive abilities. Typically, a fruit fly does not encounter incoherent motion. And there is little reason to think the ability to interpret such input would be selected for by natural selection. This extraordinary laboratory setting presents the fly with something it has never seen, and was not “designed” to see. And so, perhaps, the fly simply wiggles out. An aberrant behavior was observed, but aberrant behavior was to be expected in such a setting in the first place.

Two considerations weigh against this alternative interpretation. First, the aberrance of the behavior is fairly systematic, in that it was the dominant response of the flies. One might think that, given no clear way of handling this kind of sensory input, the flies’ behavior would default to a random set of different motor routines—a confused mish-mash of flapping. Or, perhaps the fly would simply do nothing at all. But that’s not what was observed. Rather, most of the flies displayed the same, multi-stable behavior in response to the stimulus. And even if this behavior *is* just an instance of the flies’ wiggling out, the wiggling out itself would call for explanation: We’d want to
know why so many of themwigged out in just that way. I see no way of articulating a “wigging out” proposal that gives a straightforward explanation of this. Second, if this objection were good, it would apply straightforwardly to nearly all of vision research. In most vision experiments, participants look at completely novel stimuli computer screens. Yes, at times these stimuli exploit cases in which the visual system of humans seems to go a bit haywire, as in illusion manipulations. But such cases are just the sort used to advocate for the presence of visual representation in humans.35

If all this is right—if fruit flies visually represent motion as such—there are at least three lessons we can take from this. The first is that mental representation of motion appears in far-apart clades of Animalia. It is shared by an invertebrate, most mammals, and several avian species (I discuss examples a bit more below). We cannot say with confidence on the basis of this kind of experiment whether, for example, guppies or locusts share such representations with us, rather than merely detecting motion, as the light on a driveway does. But I take it that the presence of such representation in fruit flies should dispel some prejudice against attributing mentality to animals that are morphologically very different from us. Some apparently simple creatures visually represent. And if the appealing claim that representation is a starting point for mentality is true, then this means mentality may stretch very widely across species.

The second lesson is that the criterion of multi-stability can be re-applied to other animals, and for other perceptual features and modalities, to test for the presence of perceptual

---

35 One could of course, as Gibson did, press the case further here, and insist that the research on human vision too, is similarly flawed. I dig in my heels at this point. The study of human vision is probably the most mature piece of human psychology, and I take the fact that such a well-developed and predictive science deploys a practice are strong prima facie reasons to trust them. The more developed and predictive the science, the higher the bar should be in showing that it is deeply methodologically flawed.
representation. If there is a technique that allows one to test for perceptual representations in fruit flies, surely there will be techniques that allow us to test other relatively simple animals, and more complex creatures, as well. One can imagine applying this test to virtually as many animals and perceptible features as we have the capacity and ingenuity to test. To my knowledge, this important tool has largely been ignored by philosophers of mind.

The third lesson is that in at least one case, the capacity for mental representation looks to come apart from possession of a language of thought. We have no reason to suppose that *Drosophila* can combine any of the visual representations in its repertoire. In fact, for all the evidence we have at present, motion might be the *only* property this creature is capable of representing (though this seems somewhat unlikely). Given the simplicity of its nervous systems, we ought to assume that the computations it implements are as minimal as the evidence allows. Such a minimal system might appeal to associative or Bayesian processes over simple (uncombined) representations. Though, on balance, we must accept that fruit flies have mental representations, we may assume they lack a LoT.

A few things follow from this last point. First off, if we are right to assume fruit flies lack a LoT, fruit flies are (unsurprisingly) incapable of deductive inference, and probably of object representation. By assumption, fruit flies cannot combine any representations with quantifiers or connectives. Deduction requires combination; object representation is facilitated by quantification. It might also be accomplished, as mentioned in Chapter Two, by demonstrative reference. But this would seem to call for a THAT representation which, for the same parsimony considerations, cannot just be assumed to be present in the fruit fly repertoire (and even if it *does* occur, it would be a so-called “bare” demonstrative—it could not be called on to combine with predicates to form complex demonstratives like *THAT RED DOWNWARD MOVING THING*). Notably, however, the lack of a LoT
does not rule out learning: In fact, fruit flies are capable of both classical (Pitman et al. 2009) and operant (Brembs 2011) conditioning. They might achieve this by means of associations, not between external stimuli and motor commands, but between mental representations and motor commands between representations and behavioral outputs. This kind of a mind, which pairs associations with simple (non-combinatorial) representations, offers a minimal case of representational minds in the absence of a LoT.

In the next section, I consider other species that combine representations into complex strings, and so possess a bona fide LoT. Before continuing, though, I conclude this section with more discussion of multi-stability as a mark of the mental. Multi-stability is most salient vision, but it can be generated in other modalities. For example, there exist multi-stable auditory stimuli for human beings: auditory stimuli that are ambiguous between similar phonemes. This means that the method of using multi-stability to detect mental representation can be expanded outside of vision. That is especially crucial for species for whom vision is not a dominant sense modality. Of course, there may be cases in which no test is forthcoming. I know of no multi-stable olfactory stimuli for humans, for example. Perhaps some could be generated for animals with more spatially sensitive olfaction than ours, but I am not aware of any that exist. So there may be some creatures, or even entire sense modalities, for which we cannot generate positive evidence of perceptual representation using this method.

Moreover, this method supplements others. For example, illusion is typical of perceptual representation, and so can be used as a diagnostic tool in much the same way that multistability can. And there is similar, but more widespread evidence of visual illusion in animals phylogenetically farther removed from us than macaques (Agrillo et al. 2015) or domestic dogs (Byosiere et al. 2017).
Red-tail splitfins appear to be susceptible to both the Müller-Lyer (Sovrano et al. 2016) and the Ebbinghaus illusions (Sovrano et al. 2014); bantams (Nakamura et al. 2013)—and even four-day old chicks (Salva et al. 2013)) and gray bamboo sharks (Fuss & Schluessel 2017) have also been shown to be susceptible to the Ebbinghaus. Honeybees appear to be susceptible to a version of the Kanisza triangle illusion (van Hateren et al. 1990). In a related vein, visual effects that appear to manipulate extant visual representations, such as effects of attention (inhibition of return, in the archer fish, Gabay 2013), or amodal completion (in several mammalian species, reviewed in Fujita 2006; fish species, Sovrano 2008) can provide indirect evidence. These are just a few examples among many, and stem only from the minuscule percentage of animal species that have actually been tested using methods from vision science. Likely it will prove easier to test for the presence of some markers of perceptual representation in non-human animals than others. So no one method will do the trick.

Nonetheless, using multi-stability as a mark of the perceptual presents advantages over seeking out illusions familiar to us humans in non-human animals. Traditionally, we have isolated visual illusions mostly by their effects on conscious human perception. In most cases, illusory stimuli look a certain way to us consciously. We come to understand that the way those stimuli look is at odds, in some way, with the way the world is. By and large, illusory stimuli are recognized as such relative to our own visual systems. So, even when an animal fails to show sensitivity to a visual illusion that we are susceptible to, that does not give us any evidence that they lack mental representation. They might not be susceptible to the given illusory stimulus, but still be susceptible to other versions of the relevant illusion (macaques, chimpanzees, pigeons, horses and humans are all sensitive to versions of the Ponzo illusion (for a review, Feng et al. 2017)), or to other illusions altogether. The task of imagining stimuli that would cause illusions in a non-human sensory system of any serious

---

36 Interestingly, despite extensive tests, the evidence is still mixed for pigeons (review in Lazareva 2017)
complexity, but not in our own, is a tall order. By contrast, multi-stable perception is less
demanding. One need know, generally, only which features of the perceptible world a creature is
sensitive to. Then one has some idea of what kind of stimuli might plausibly generate multi-stable
percepts: those that combine two or more of those features in a potentially ambiguous way. One can
imagine, for example, generating bi-stable stimuli for the echolocation systems of bats, for which a
precise psychophysics is now emerging (e.g. Warnecke & Simmons 2016).

Interestingly, this epistemic advantage over evidence for susceptibility to visual illusion does
not stretch, even mutatis mutandis, to modal or amodal perception. Those visual capacities are
individuated by success at the environmental challenges they help address. If a creature sees an
orange as partially occluded by the apple, rather than as having part of an apple-shape cut out of it,
that will be, all else equal, evident in how the creature behaves toward the orange. So testing for this
capacity is relatively straightforward. On the modest assumption that no creature would visually
complete object representations but act in a way that ignores that completion, failure at such tasks
should be positive evidence of a lack of the capacity for completion. One can go further: Since
completion is such a rudimentary feature of visual representation, failure in such tasks is (defeasible)
positive evidence for lack of representation tout court. While a representational visual system that
can be duped easily by occlusion is conceivable, navigation, foraging, or hunting by means of such a
system would be useless in most natural environments. Evidence for modal and amodal completion
are comparably useful to evidence for multi-stability, while evidence for susceptibility for visual
illusion is not.

So far, I have considered some reasons for thinking that surprisingly simple creatures might
possess minds, in the sense that they can mentally represent. And I’ve argued that there is some
reason for supposing that such minds might, nonetheless, not be combinatorial. In the next section, I consider cases of minds on the other side of the combinatoriality gulf: those with LoTs.

Three Arthropod Languages of Thought

In this section, I describe evidence for three arthropod LoTs. The kinds of LoT under consideration may be minimal. For all the evidence suggests, they may not possess the full cluster of traits often associated with a LoT. Nonetheless, they are collections of mental representations, at least some of which can be combined to form complex strings. Empirical evidence from several arthropod species supports that syntactic combination. I consider some attractive non-combinatorial explanations, and show they fail to explain the data adequately.

We can start by presenting three more pieces of empirical evidence. I will describe each briefly, then turn to an extended discussion of all three afterwards. The first piece of evidence is a recent study from Lars Chittka’s lab documenting flexible behavior in bumblebee social learning (Loukola et al 2017). The best explanation of the flexibility is a fairly abstract kind of shape representation. The experimenters trained nine bumblebees to push one of three balls on a platform (the other two balls were glued down) to a central well in order to access a sucrose reward. The ball that this group of bumblebees were trained to push was yellow, and it was the farthest ball on the platform from the central well. Then, the experimenters allowed other bees to watch the trained bees perform the task. When placed on a similar platform later, those “observer” bees then were able to perform the task themselves. (This movement—pushing a round object from behind—is unusual for bumblebees; the authors report that it has not been observed in other contexts). Interestingly,
the bees that learned by observing their conspecifics acquired the sucrose more quickly, and in larger numbers, than did bees learned in other ways: Some had been shown a ball moving to the target area without being pushed, and others had been shown an artificial “bumblebee” on a stick pushing the ball. Yet what is really striking about this study, for our purposes, is not merely that the observer bees engaged in social learning. Rather, it is the other the manipulations of the study. The platform used by the first round of bees—the ones that were observed—differed from the platform used subsequently by the observers. The first group of bees was trained to push just one, yellow ball from the edge of the enclosure into the target area containing the well. The other two balls were glued down, so the observers only saw them push that ball into the well. Nonetheless, a majority of the observer bees did not go for the ball which, like the one they had observed being pushed, was at the edge of the enclosure. Rather, they went for the ball that was closest to the target area. They did this even when the closest ball was black, not yellow—and despite the fact that the farther balls were yellow.
It is tempting to conclude from this that the bumblebees formed an abstract goal: *Push a ball into the well*. Then, they compared different behavioral options that could be categorized as pushing a ball into the well, and selected the least costly. A bit of reverse engineering can help flesh out this way of interpreting the result. Suppose one wanted to build a decision-theoretic machine that could put a ball in a well with the efficiency of the bumblebees. Such a machine would need a fairly general way of characterizing instances of balls in that well (and not just yellow or far away balls). And it would need a means of selecting actions apt to that goal. In most cases, the machine will have a range of physical routines it can perform that result in the goal state: It might push from one piece of its apparatus, or another; it might use a circuitous route, or a direct one. In most cases, the action
or actions that minimize effort will be optimal. The machine would thus need to weigh its options over a hypothesis space.

The second piece of evidence comes from work on two kinds of Foundress, or paper wasp. It suggests paper wasps are capable of transitive inference. That is a kind of reasoning that draws a conclusion about a relation between two items that have not been compared before. Typically, they take the form: aRb, and also bRe; therefore aRb. A familiar sort of transitive inference might be “I am less skilled at hammering than Noah, who is less skilled at hammering than Jelscha; I am less skilled at hammering than Jelscha. The primary interest of transitive inference in paper wasps (or any species) is that in order to accomplish this reasoning pattern, the animals need to form preference orderings over which such inferences may occur. Most plausibly, these are LoT strings.

The experimenters trained paper wasps to associate colored patches to electric shocks. They presented two patches at a time, one on each end of the wasp’s enclosure. Each wasp began each trial in the center of the enclosure, hemmed in by two plexiglass walls. The floor of much of the enclosure, including the area hemmed in by the plexiglass, was electrified. But in each trial, one part of the enclosure, which was marked by one of the colored patches, was not electrified. Thus, when the plexiglass walls were removed, the wasps could find reprieve from the shock by moving toward one of the two patches. Three of the five patches appeared as often on the electrified side as on the non-electrified side (which varied left to right). One patch only appeared on the electrified side. And one only appeared on the non-electrified side. Letting A, B, C, D, and E name five color patches\textsuperscript{37}, we can describe the training in the following way:

\textsuperscript{37} NB: The hierarchy positions of colors were counterbalanced across wasps. The letters A-E name positions in the hierarchy, not particular colors. Each color patch figured as A, B, C, D, and E, depending on the wasp.
$A_0B-, B_0C-, C_0D-, D_0E-$ ($0 =$ no shock, $- =$ electric shock)

As a result, after training, the wasps were sensitive to *pairings* of patches. They would, e.g., move toward A when it was presented along with B, but they would move toward B when it was presented with C.

At test, the wasps were presented with a novel pair of panels, BD. Both B and D had been presented to each wasp, with and without shock. But the pair had never been presented together. The wasps more often moved toward B. This suggests that they were sensitive, in some way, to the hierarchical nature of their training. The tempting conclusion is that formed a hierarchical representation with (roughly) the form: $A>B>C>D>E$. Then they performed a transitive inference: $B>C>D$; therefore, $B>D$

*Figure 3: Polistes wasps learned associations of colors (here: indigo) and shock when those colors appeared with others (here: violet).* 

*Reproduced from (Tibbetts et al, 2019), Supplementary Materials*
The third piece of evidence focuses on a darling of animal cognition research: the honeybee. Somewhat famously, honeybees that have recently come across a high-quality source of nectar will perform a characteristic set of movements when they get back to their hives. Von Frisch (1967) noticed that the dances vary along predictable axes. These correspond to the quality of the nectar, as well as its distance and direction from the hive, relative to the sun. Interestingly, not all bees who are in a position to witness the dance leave for the nectar, even after witnessing more waggle runs than is usually required to find the nectar. A predictor of whether a bee will head to the nectar is whether the bee has discovered a preferable source of nectar (Grüter et al. 2008; Grüter & Farina 2009). This flexible behavior is easily explained by attributing to the bees (a) representations of nectar-quality and (b) some decision-making process that operates on those representations. (Note the similarity to what we have proposed for the bumblebees). It is tempting to attribute something even stronger than (a) to the honeybees: a complex representation of (roughly) the form NECTAR SOURCE 1 > NECTAR SOURCE 2.

Associationist models would seem to be a plausible alternative to the LoT-friendly conclusions sketched above. But the most straightforward versions of such models struggle to account for these results. And more complicated associative explanations do not explain them as well as the inferential explanations just canvassed.

In the bumblebee case, the clearest alternative is that bees are associating the balls they see, (more specifically, their hivemates’ pushing those balls) with the sucrose reward. It is difficult to explain how any mimetic social learning can be accomplished by associative learning, since the
observer is most plausibly associating another’s actions with the outcome, not their own (for a nice discussion of this problem, see Heyes & Papineau 2006). But, assuming that a clever associative account of social learning of this kind is possible, it is crucial that the bees in this study rely on shape, *but not on color*, in action selection. Why, on any associationist model, should we expect such a bias in favor of shape? Moreover, the bumblebees showed a preference for the *closest* ball to the reservoir, though they had observed a farther one being pushed. If the social learning in question is largely or mostly grounded in associative learning processes, this is mysterious. Meanwhile, if the animal possesses the fairly abstract representation BALL, describing what was learned is fairly uncomplicated: The bees learned that putting *balls* into the well leads to reward.

A fallback associationist strategy is to say that BALL itself is associated with the reward. This strategy admits that there is a fairly abstract shape representation guiding the bees’ behavior, but denies that it is combined with other representations. This strategy can accommodate the fact that the bumblebees are selective about which of the stimulus ball’s properties are relevant to obtaining the reward, because it is compatible with bees’ representing shape as such. It does not explain how or why the bees take the fact that the objects are balls to be relevant (since pushing objects in the way shown is a very unusual activity for bumblebees, roundness as a cue for mobility may well be irrelevant to them). But it does not rule such a preference out.

However, this strategy struggles for a different reason. BALL must somehow interface with a representation of the target area, and some representation of pushing, to yield the action. It simply cannot be, for example, that BALL, the target area, and the pushing, are each independently

---

38 As Tyler Burge (2014) has pointed out, there may be generic, but still pictorial, representations of shape. So the representation need not be BALL. It could be a generic pictorial representation of a ball. Either will do for the purposes of my argument.
associated with reward. Such an associative explanation cannot explain why the bees performed actions that were not reinforced, but that were nonetheless decision theoretically optimal. Moreover, if it were true, we would have expected the bees should push the wall of the enclosure, or simply walk into the well, which they did not do.

Still, there is a higher-order associative explanation in the wings. In higher-order association, learned associations are themselves associated. This can allow for individuals to form associations that include complex configurations of cues. Perhaps the bumblebees are engaging in learning of this kind, in which the configuration associated with reward involves a ball, pushing, and the well. Curiously enough, this strategy falls on the same sword that the simplest strategy did. On the face of it, it is unclear why the configuration should not include associations with yellow balls, rather than balls simpliciter. The familiar problem of explaining how BALL interfaces with the rest of the configuration arises for these higher-order associations. There may be strategies available to associationist explanation at this point, but they begin to look unbecomingly ornate.

As for the wasps, the trouble for associationist explanation lies in the fact that the wasps have never encountered the pair BD before. Since both B and D have been paired with shocks, associationist accounts of learning should make no prediction about behavior here. The wasps should navigate toward B or D at chance. They do not: They preferentially walk toward B. If the wasps are building a hierarchical representation of these items, however, this is easy to explain. It is B’s and D’s relative places in the hierarchy that explain the behavior. This representation might take the form of a map of hierarchy relations (as noted earlier, Camp 2009 argues for a view of this kind to explain baboon hierarchy representations), or a sentence in the wasp’s LoT.
There is at least one well-developed alternative explanation for the paper wasp case in the literature. Some have advocated for a different interpretation of five-item transitive inference tests, like the one the wasps succeeded at. This alternative is based on Value Transfer Theory (Fersen, Wynne, Delius, and Staddon 1991; Allen 2006). The researchers reasonably point out that A was never associated with shock. And E was always associated with a shock. So E’s presentation with D might have led to a transference of negative valence from E to D, respectively. But if that happened, it should be unsurprising that B was preferred over D. While B was neutral, D had been transferred a negative valence.

But this kind of alternative takes on strong assumptions that transitive inferential explanations need not. First it is unclear why valence should transfer from E to D. For one thing, D is neutral when presented with E, and it is precisely this pattern, of D being neutral when presented with E, that the wasp learns to recognize. If anything is transferred from that learning to test, why shouldn’t it be that D is (completely) neutral? This is bolstered by the fact that value transfer has been demonstrated in associative learning, it has only been shown to occur for items that are presented contemporaneous. It has not been shown to hold across separate trials (Vasconcelos 2008). This matters because at test, E is absent. It might well be that value transfer only holds from E to D when E is presented to the organism along with D. This seems reason to default to an explanation of the learning that involves preference ordering of some kind.

In the honeybee example, the case against the associationist is somewhat weaker, but it rules out the most straightforward associationist explanations. The kind of explanation I have suggested, which relies on a GREATER THAN operator, does not fare better than some more complex, but still plausible, alternatives: pairings of magnitudes or valences with maps, or preference orderings as lists. To rule out these more precise alternatives, we will need to appeal to a different set of observations, which suggest that we already need to appeal to a LoT to explain honeybee decision-making. This thereby strengthens the case for LoTs in this case too. I describe that evidence in Section III.
We can start by ruling out two simple associationist stories. Operant conditioning can explain why a bee that perceives a waggle dance successfully navigates to the indicated nectar source. Changes in navigation in response to specific dance movements lead to reward. So it might seem like this is a reasonable explanation of behavior of the honeybees. But operant conditioning cannot explain why a bee who has perceived the dance will return instead to its preferred source. The behavior of the bee depends on the remembered quality of some nectar other than the one indicated by the dance. It is unclear why this should interfere with an association previously formed on the basis of a dance-navigation pair. A more sophisticated theorist might take issue with the simplicity of the associationist story sketched here. She might, as with the bumblebees, appeal to second-order associations to explain the behavior. But it is not clear how such complications of the story address the problem. The challenge is to spell out how second-order (or higher) associations could explain why the honeybees behave in the ways described. It cannot be, for example, that the bees learn to ignore the waggle dance if they have found another source of nectar, because (a) viewing the waggle dance seems to induce the departure for nectar, and (b) the bees’ navigation is sensitive to whether the nectar source they found previously is of higher quality. Meanwhile, a strategy that posits combinations of representations can help itself to simple explanations. On such a strategy, the behavior needn’t be mysterious, nor the explanation baroque.

Another move one can take is to modify, not the kind of association at work, but the way that the nectar sources are represented. Suppose honeybees possess an analog magnitude means of representing overall nectar quality. The abstract property of nectar quality might be represented in the same continuous, mass-like way that children, human adults (at times), and many animals represent quantity. It need only associate such representations with representations of locations (for
example, in a cognitive map (Gould 1986). The animal might then adhere to a rule according to which, once it has been prompted to seek nectar, it should always navigate to the location associated with highest quality nectar source. Alternatively, the locations might themselves be directly associated with positive valences of various strengths, so that, once prompted to consult its map of nectar sources, it selects the destination with the highest associated valence. In either case, it looks as though one may be able to avoid attributing to the bee any combinatorial syntax. All we need, instead, are associations of analogue magnitude representations or valences, and some means of representing at least two different locations.

This kind of strategy has the virtue of simplicity, because it is not such a stretch to imagine such structures exist in the honeybee mind. Analogue magnitudes are widespread in the animal world. If bees possess representations of quantity at all, they will possess analogue magnitude representations. And valence is a common component of associationist explanation generally. Finally, cognitive maps in honeybees, while somewhat controversial, are a fixture in the comparative psychological and behavioral ethological literature. They have been extensively discussed and tested on bees for several decades now. By comparison, representations with combinatorial syntactic structure, such as \texttt{NECTAR SOURCE 1 > NECTAR SOURCE 2}, are a fresh ontological commitment.

The case for a LoT-style explanation here could be strengthened by independent evidence that honeybees can combine representations syntactically. Ideally, this would be evidence for representation of \textit{ordering}. Such evidence has emerged recently in a series of comparative

\footnote{It might even associated with simple motor commands that would navigate the animal to those locations, or some to some other navigational apparatus the honeybee uses, though these alternatives appear explanatorily inadequate for the honeybee (for an interesting recent discussion, see Rey and Knoll 2017).}
psychological studies of honeybees, in Section III. The pairing of this evidence with the arguments in Section III yield a larger argument for preference orderings of nectar sources in honeybees.

Before continuing to Section III, however, I consider an alternative to sentence-like strings for representing preference orderings. Might honeybees be building lists of nectar sources, and ordering them by preference? This strategy does not call for an ordering operator such as GREATER THAN. It thus solves the challenge I raised but does result in a sentence with a subject-predicate form. Any string with one or more GREATER THAN operators can be redescribed as a list. All else equal, lists are as explanatorily powerful as a GREATER THAN operator.

Though it is a viable alternative, lists are compatible with the claim that honeybees possess a LoT. While it is one thing to say that the computation can be effected without appeal to an explicit GREATER THAN operator; it is another to say that it can be effected without a LoT. Consider

[List 1:

NECTAR SOURCE 1
NECTAR SOURCE 2

and List 2:

NECTAR SOURCE 2
NECTAR SOURCE 1.]
Both contain the same representations, ordered differently. But why does top-to-bottom order, rather than bottom-to-top order (typically) correspond to preference order in preference orderings? Because there is a rule that specifies how constituents of the list may be combined. This may be a matter of convention when we are cataloguing Wimbledon winners or albums of the year. Perhaps, had history been different, our lists might have mostly read bottom-to-top (and perhaps, in some cultures, they are read that way, though I do not know of any). But if lists explain the behavior of the honeybees, this cannot be a matter of convention. There must be some rule that governs or describes how those lists are composed, and how they figure in the computations that lead honeybees to select a nectar-source. It is difficult to see a non-ad-hoc reason for calling this something other than a syntactic rule. On the characterization of LoTs we have been relying on all along, such a list qualifies as a LoT string

It is worth taking stock at this point. First off, I have provided evidence that bumblebees and paper wasps possess a LoT, and that this interpretation of the evidence is preferable to the most plausible alternatives. As far as this evidence goes, it bears noting that this is all I have argued for. A LoT might possess some forms of combinatoriality, but lack others. And so it is possible that the arthropods under consideration possess only “incompletely” combinatorial minds. For example, perhaps the paper wasps described above possess a GREATER THAN operator that helps them decide how best to navigate away from electric shocks. But how do they represent A, B, C, D, and E? The paper-wasp LoT might lack sentential combinatoriality: It might contain no predicate-argument

---

40One might take issue with the characterization of LoTs I have been relying on, of course. One might wish, as a desideratum of any version of the LoTH, for LoTs to be characterized so that lists are not LoT strings. But whatever one’s account of LoTs turns out to be, it cannot be denied that lists of this kind will be quite similar to LoT strings involving ordering operators, for the reasons outlined above. As I suggested at the close of Chapter 2, considerations like this might motivate some to delineate a new representational genus of which LoTs are species: combinatorial-representational systems.
structure. The representations A, B, C, D, and E might be one-word sentences, picture-like representations, or cognitive maps. The greater-than relation might be represented with an operator that combines with such representations, rather than a predicate that combines with pairs of arguments. Second, I have provided evidence that honeybees are able to integrate information from disparate sources in nectar foraging. This is evidence for an impressive degree of representational power and flexibility in the honeybee mind. But on its own, it does not necessitate a LoT: It is compatible with the claim that honeybees possess cognitive maps, and that they pair parts or regions of those maps with valences or magnitudes. Additional evidence is needed to show that honeybees possess combinatorial minds.

Number in the Honeybee LoT

In this section, I argue that honeybees represent number in a way that is best explained attributing to them symbols in a LoT. I survey four recent comparative studies from the Center for Research on Animal Cognition, in Toulouse, France. Jointly, they suggest that honeybees can pair mental representations with discrete quantities, such as 2 or 4; they can represent number in ways incompatible with other known means of representing quantity: analogue magnitudes and subitizing; and that they can combine representations of quantity in order to select a source of sugar.

It is, of course, of inherent interest to know if, and how, honeybees and other animals can represent quantities in their environment. This is a project for anyone interested in animal cognition. But in the context of the present discussion, the latter question will be of special interest to us. Some ways of representing quantities are especially congenial to explanation by a LoT.
There are at least three ways that a mind might represent numerosity. Number representations, such as ONE or FOUR, function to represent discrete numbers. This is, most straightforwardly, the way one might imagine a LoT representing number: by means of unique symbols for numerosities. They can be contrasted with at least two other means of quantifying amount: analogue magnitude representations and subitizing.

Magnitudes are not discrete, but continuous. Thus, representations that are formatted as magnitudes are isomorphic to the quantities they represent. Literally, physical features of the brain, such as neural firing rates, scale up or down as a function of increase or decrease in represented quantity. This feature of magnitudes manifests itself behaviorally in a psychometric law known as Weber’s law: Our ability to discriminate by means of two magnitudes is a matter of proportional, rather than absolute, differences. As quantities increase, so does the brute amount of difference it would take for us to distinguish them. For consider the asymmetry between figures 4 and 5.

Figure 4: Which of these clusters is larger in number?
It is fairly easy to tell, without counting, that one cluster in Fig. 4 is larger than the other. And, by contrast, it is difficult to tell, without counting, which cluster is larger in Fig. 5.

Figure 5: Now, which of these clusters is larger in number?

But both clusters differ by the same amount: two circles. It is the proportional difference that changes from Fig. 4 to Fig. 5: 14/16 is greater than 6/8.

Subitizing is a rapid, seemingly effortless extraction of numerosity for a small number of objects. Subitizing is presumably what makes it so effortless to tell, without counting, exactly how many objects are in each of the clusters below:
The upper bound on subitizing, in humans, is the same as a commonly assumed (item-based) limit on working memory: 4. Subitizing is typically taken to be enabled by forming object representations, which are then stored in working memory. For example, to recognize the cluster at the upper left has three items, I might token three visual representations of circles, then transfer each of those items to working memory. Since it is the individual objects that are represented, and not a quantity or numerosity, the number of objects in a scene is represented only implicitly by subitizing. It is likely to be rapidly inferred from the number of items stored in working memory.

Adult humans possess both of these kinds of representation, but they also appear to possess a third. Many learn and manipulate discrete number symbols, such as ‘3’ or ‘64’. And adults can think thoughts about zero and negatives which are impossible to represent by means of a magnitude or by subitizing (I discuss reasons for this below). So it seems that adult humans possess some other means of representing number, by means of discrete symbols, such as THREE, SIXTY FOUR, or ZERO. These abilities emerge latest in development, and it has been argued that the capacity for symbolic representation of number piggybacks (in some way) on these more ancient capacities (Gallistel & Gelman 2000; Mandelbaum 2013b). But that it is grounded in those capacities does not rule out that it exists, and that applications of it allow for things that are not possible by means of those systems alone—for example, calculations involving zero.
Recent studies have probed whether honeybees might possess (or learn) such symbols for numbers. The results are instructive. Howard et. al (2019a) trained honeybees to associate numbers of shapes with arbitrary symbols (an ‘N’ or an upside-down ‘T’). The bees flew into a Y-shaped maze with symbols over the entrance of each arm of the maze (see figure below). They were rewarded with sucrose solution if they chose to go down a tunnel bearing a sign with either two or three objects on it (depending on the condition). They would be ‘punished’ with quinine solution if they chose a tunnel bearing a sign with the other number of objects on it (three or two). Thus, it looks as though the bees learned to pair particular symbols with particular numbers of shapes, because those pairings were predictive of reward (while other pairings were predictive of punishment). To rule out the possibility that the bees were relying on low-level features of the cards they had been trained on, rather than the number of shapes depicted on them, the authors tested the same bees again. This time, they used shapes, colors, sizes, and locations for the items that were different from those presented initially. The honeybees tested were able to transfer what they had learned the first time around to this task. It appears they really paired the number of shapes on the cards with the abstract symbols (they may have also paired those low-level features, but they did not pair only those low-level features). A tempting interpretation here is to say that honeybees possess, or have formed, representations of those numbers. They represent number by means of symbols such as TWO or THREE. Moreover, they can associate novel symbols with those LoT representations.
Figure 7: Honeybees using signs in a maze to navigate to a sucrose reward. The bees learned to associate pairings of numerosities of elements (e.g. 2) and arbitrary symbol (e.g. ‘T’) with reward. Even when they only saw the items sequentially, they were able to exploit this association to locate the reward in a maze. Reproduced from Howard 2019a.

This does not show that the bees are combining those representations of numerosity to form new representations, or that they are performing logical inferences on the basis of them. There are reasonable non-inferential explanations of the behavior: These appeal to associations of sequences of symbols with reward: the ‘N’ or ‘T’, followed by honeybee number representations (TWO or THREE). Notably, this picture still adverts to the kinds of representations that I argued they likely possess:
TWO and THREE. However, it does not involve any *combination* of those representations with others, or transformations of representations including them.

However, another aspect of the results of this study places a constraint on associationist pictures of this type. The bees failed at reversal tasks: Bees trained on symbol-numerosity associations (i.e., tasks in which the symbol was on the outside of the maze, and the panel with shapes on it was inside the maze) were tested on numerosity-symbol associations (shapes on the outside, symbol on the inside)—and vice-versa. In this test, the bees were not better than chance at selecting the arm of the ‘Y’ containing sucrose. This datum is more difficult for the associationist to explain. In fact, it is at odds with what Quilty-Dunn & Mandelbaum (not yet published) argue is a central feature of associationism: Generally, if X is associated with Y, Y is associated with X. Generally, such reversal tasks should be a piece of cake if the symbols are associated with number representations. Similarly, the more one sees asymmetries in the relations an organism learns, the less likely it is that associations explain the learning. That said, there may be a more complicated associationist explanation at work. The associations would need to be sensitive, not just to the tokening of mental representations and the form of written symbols, but also to the order of occurrence of those items.

By contrast, there is no comparable constraint placed on explanations that rely on combinations of representations. Nothing about those explanations predicts the observed asymmetry in learning. But it is also fairly easy to accommodate them: Plenty of syntactic relations are asymmetric. It is not that case that if I rob you, you rob me; nor that if P > Q, Q > P. So, while there are plenty of options available to the associationist, this is one less thing to worry about for the advocate of honeybee LoTs.
Another study, from the same lab, adds support to the that honeybees have discrete number representations. Bortot et al. (2019) presented honeybees with a similar Y-shaped maze. Once inside the maze, the bees could again choose to go down either arm, and again, each arm was marked with a different number of shapes. All the bees were trained to go down an arm of the maze marked with three shapes, the other arm being marked with two objects, or four. As a result, half the bees were trained to go down an arm marked with a larger relative amount, and half to go down an arm marked with a smaller relative amount. At test, experimenters presented the bees either with the same choice as in training, (with the objects differently organized and shaped, to control for low-level features), or with the opposite choice. That is, if a bee was trained on 3 vs 2, it could be tested on 3 vs 2 again, or on 3 vs 4. In both conditions, the bees opted, at a rate significantly higher than chance (about 60%), to go down the arm marked with three objects. The first result suggests that the learning was not based on low-level features of the training stimulus. The second suggests that they did not (only) learn to opt for the larger of the two quantities. Together, they suggest that the fact that the arms were marked *three shapes* was guiding the successful navigation toward sucrose.

![Figure 8: Honeybees learned to associate signs with three items on them with a sucrose reward. They selected cards with three items on them even when the authors varied the low-level features of the items on the card, and when they varied whether three was fewer or greater than the number of items on the alternative card. Reproduced from Bortot et al. 2019.](image)
Here, as before, nothing rules out associations of number representations to reward. The bees may even have been relying on a simpler strategy than the one proposed above: one that pairs discrete THREE directly with reward.

But in a related study, Howard et al. (2019b) provide evidence that is very difficult for even the sophisticated associationist picture sketched above to accommodate: Honeybees seem to be sensitive to the fact that zero objects is fewer than some objects.

It is worth spelling out, before looking at the result, the two ways in which this evidence is significant. The first thing to note is that the ability to represent zero at all strongly suggests the presence of a means of representing quantity that is neither a magnitude nor subitizing. It is not obvious how magnitudes could represent zero: Since magnitudes are isomorphic in scale to their representata, to accurately represent a null quantity would require a magnitude that has a null “magnitude”—i.e., for there to be no magnitude. And similarly, one cannot subitize to represent zero: An “accurate” representation of zero objects by means of object representations would include zero object representations. It would mean not subitzing at all.

The second, and more significant, thing to notice is that if bees are cardinally ordering zero, their zero-representation very likely combines with others. If the bees are (somehow) sorting items on the basis of a less-than or a greater-than relation. It is plausible to assume that whatever computation instantiates this comparison is using representations of a similar format. There is independent evidence, from human development, for this assumption. That children can compare amounts fewer than four, such as 1 and 3, and can compare amounts greater than four, such as 12 and 14, but fail
to compare amounts that cross these categories: for example, 3 and 6, or 4 and 8 (Feigenson et al., 2002; Xu, 2003; Lipton & Spelke 2004; Feigenson & Carey 2005; with an exception for ratios greater than 4:1, Cordes & Brannon 2009). The typical explanation of this is that at that stage of development, they represent amounts fewer than four by subitizing, and amounts greater than four by analogue magnitudes. The problem for them would then appear to be an inability to compare amounts represented in distinct formats. Either bees can do something that human children cannot do—namely, compare quantities in distinct formats—or they have some common code in which to make comparisons of amounts. If the latter, the code cannot be magnitudes or object representations, because the comparison involves zero. So if the bees can perform such comparisons, that would bring us right up against the conclusion that the bees are performing numerical comparisons in a LoT.

Now we can turn to that evidence. The honeybees, yet again presented with a Y-shaped maze, were trained on dozens of sequences of pairs of white squares. The squares had 1-4 black objects on them, varying in configuration, size and shape. The honeybees learned to select the square in each pair with the greater or fewer number of shapes. For example, when presented with a card that had four objects on it, and a card that had three objects, bees trained on the “greater than” relation learned to select the card with four shapes, and bees trained on the “fewer than” relation learned to select the card with three shapes. After training, the bees saw novel comparisons, including cards they had never seen before. For example: a comparison of three and five. The crucial test was a novel comparison that included a blank card, i.e. one that that depicted zero objects. The bees continued their pattern for these novel pairs, including for the pairs involving zero. Bees trained on the “fewer-than” relation selected the empty card, and bees trained on the “greater-than” relation selected the non-empty card (even if that card contained just one depicted element!).
Again, there are two lessons to draw here. First, because it is difficult to understand how a mind might achieve such comparisons by means of magnitudes or subitizing, this offers impressive supplemental evidence for the claim that honeybees represent numerosities. Second, and more importantly, this offers evidence that the comparisons may be happening in a LoT. Either honeybees do something young children cannot do—compare quantities represented in distinct formats—or they are making comparisons in a common format. Since that comparison includes zero, it cannot be a comparison of magnitudes or object representations, so it is likely a comparison of symbolic representations of number. But it is just very difficult to imagine what such a comparison might look like if it did not include some means of representing the cardinal ordering of the symbols, such as a GREATER THAN symbol. While this is not decisive evidence in favor of such
combinations in honeybees, the alternative to attributing such combinations to them would be just as astounding.

In addition to these two points, there is also more speculative suggestion we can draw about the syntax of honeybee representations. If the common-code interpretation is correct, this might suggest there is sub-sentential structure in the honeybee mind. Very plausibly, the representations of numerosity that explain this result are concepts. Just as it is implausible that zero might be represented by means of magnitudes or object representations, it is equally strange to suppose that it might be represented by means of vectors, maps, or pictures. One-word sentences might seem an alternative here. Since they can represent quantities at whatever level of abstraction one likes, they are not hindered the fact that they might represent nothing. But which sentences would do the job? The most plausible option appears to be a symbol which expresses the negation of an existentially quantified expression, such as THERE-ARE-NO-OBJECTS-HERE. But it is not obvious how this might combine with other representations in a way that represents a cardinal ordering. It is not obvious, for example, that much sense can be made of the following: THERE-ARE-NO-OBJECTS-HERE < THERE-IS-ONE-OBJECT-HERE. By contrast, it is not difficult at all to see how a creature might represent a cardinal ordering by means of sentential combination of concepts. Perhaps, for example, the honeybees rely on a GREATER THAN operator like the one mentioned earlier, to build: ZERO IS LESS THAN ONE. This is meant more as a challenge to those who would deny that the relevant representations are concepts than a defense of the claim that they air. That said, further evidence for the common-code interpretation of cardinal ordering of zero, or for combinations of number representations with other concepts (e.g., evidence for FOUR SQUARES or FOUR BLACK SHAPES) could strengthen this speculation.
Conclusion

This chapter described two lower boundaries of mentality, mental representation and combinatoriality, by focusing on arthropod psychology. Representational minds may lack combinatoriality, as we can only imagine is the case of the fruit fly. But, as the fruit fly itself demonstrates, they may be far more widespread than is commonly imagined. With the right empirical tools, including, as I have advocated, tests of multi-stable perception, we can tell just how widespread it is. Combinatorial minds, a species of representational minds, are less common than representational minds, but still clearly widespread, as the bumblebees, honeybees, and paper wasps demonstrate. Honeybees appear to possess, not just a capacity for mental representation, and for combinatoriality, but an astounding capacity to represent number symbolically, including the number zero. This is astounding in its own right. But it also strengthens the case for a honeybee LoT, since it suggests that honeybees compare number in a common code with whatever system might represents zero—and this cannot be achieved by magnitudes or subitizing. Because, as illustrated in Chapter Two, evidence for combinatoriality cannot be taken to be evidence for all forms of combinatoriality, the work of detecting combinatoriality in a given creature or other cognitive system is only the beginning of understanding its LoT. We must then get clear on the syntax of that LoT. But, as illustrated by some exciting recent work on honeybee representations of numerosity, there is room for speculation, and there may be empirical tools that can help answer this question.

The sort of LoTs I have described for bees and wasps here might seem a bit alien to the LoTs typically thought of in the context of the LoTH. That posit was proposed, first and foremost, in the context of explaining human cognition. There is some continuity between
arthropod LoTs and the putative human LoT, of course, but perhaps unsurprisingly, it looks to be minimal. For example, there is no reason to suppose that these arthropods have more than a minimal representational repertoire. They may only mentally represent a limited range of features of their environment, such as motion, objects, locations, or nectar sources. One can only imagine that they lack conceptual counterparts of sentential connectives, which might underpin a capacity for deductive inference. We cannot even say confidently which rules actually govern the combinations of their mental representations. We do not know, for example, whether any of them are recursive, or if there is a (competence-based) limit to the number of acceptable iterations of them.
Chapter Five: Some Evidence for Disjunction and Negation in Baboons, Chimpanzees, and an African Grey Parrot

In this chapter, I survey some recent results from tests on olive baboons, chimpanzees, and an African grey parrot that suggest these species possess conceptual counterparts of sentential connectives, and thus what I have called supra-sentential combinatoriality. As a result, there is reason to think these animals possess a more sophisticated LoT than the other animals so far surveyed, one that might implement a much wider range of psychological competences.

The Evidence

I start with a few words about why evidence for representing logical connectives is significant. Evidence for representation of logical connectives just is evidence of a capacity to build complex representations. If some logical connectives are represented as such, then they really must compose other representations in ways that respect rules of combination. A well-formed string including OR also includes two constituent sentences. If a mental counterpart to a logical connective does not take two sentences as constituents, it cannot be OR. But since evidence for representations of connectives is evidence for complex representations in one's thoughts, it is also evidence for a kind of combination. As such, it is some evidence for a LoT. Suppose we were to discover that organisms of species S possess a psychological equivalent of ‘and’. We now have excellent reason for supposing that Ss can token the complex representation type P AND Q—assuming, of course, that Ss are clever enough to represent strings that are three symbols long. We also have reason to suppose they can represent P AND Q AND R, and P AND Q AND R AND S, and so on—again, assuming they can
store all that in memory. This is a recursive feature of thought. An alternative possibility is that an OR-like representation might exist; this representation would be just like OR in every way, except one. According to the grammar of some animal’s thought, it can only be applied to sentences that do not already contain an OR. It would fail to be recursive. Nothing I will say here can rule this possibility out. But, as I stressed in Chapter Two, whether a LoT is combinatorial should not hinge on this question.

Of course, the fact that an animal can think P AND Q does not by itself tell us very much about the nature of its P-thoughts or its Q-thoughts. Those thoughts might be composed of atomistic parts. For example, if P stands in for ‘Jim loves James’, it might be that it is composed of constituents ‘Jim’ and ‘loves James’ (and the latter, then, of ‘loves’ and ‘James’; and so on). But it needn’t be so composed. The thought that Jim loves James might turn out to be a Jim-loves-James thought, and the ability to think it might be independent from the ability to think James-loves-Jim thoughts. In such a case, we would not have composition of constituents of simple sentences in a LoT, and so at least one sort of combination—sub-sentential combination—would be missing.

But even if the presence of connectives cannot tell us about the sub-sentential structure of thoughts, I take it to be significant progress toward identifying a kind of non-human language of thought. We will have shown that some nonhuman organisms do compose simple sentences into more complex ones. Concatenation in propositional logic is a kind of composition. P AND Q is a sentence, with proper parts P, AND, and Q.

I have been talking about conjunction so far, because it is a straightforward case. But precisely the same lesson applies for other sentential connectives: NOT, OR, ONLY IF, and IFF.
bracket for the moment the question of whether any thought really is mediated by something like the material conditional. There are plenty of reasons to think that the material conditional does not underlie conditional reasoning. But it seems plausible that a language of thought should include a logical constant with a role comparable to that of the material conditional: Even non-logicians appreciate that there is something acceptable about modus ponens. Alternatively, a creature that lacked material conditional might possess both negation and disjunction, and so potentially be capable of performing inferences equivalent to those involving material conditional (since in first-order propositional logic, \( P \supset Q \equiv \neg P \vee Q \)). In any case, the examples I consider moving forward include disjunction and negation, and so this is not terribly pressing.

I now present the evidence that I take to suggest that some non-human animals really do token representations of disjunction and, in one case, negation. A recent pair of studies report four olive baboons (Ferrigno et al., in press) and an African grey parrot, Griffin, (Pepperberg et al. 2018) all succeeded at a task most easily solved by means of disjunctive syllogism. I here discuss the baboon case at length, though the lesson to be drawn from the parrot case is nearly the same (though the parrot study involved just one participant with many years of experience as a test subject, and so is more susceptible to the criticism that the sample is unrepresentative of African grey parrots generally). The task, adapted from one Mody & Carey (2016) used on 2.5- and 3-year old children, is as follows. In an initial phase, olive baboons are familiarized over several weeks with a task in which food is placed in one of two adjacent cups, but not both. The experimenter places the food behind a screen, so that the monkey does not know which cup the food ends up in. The experimenter shows the monkey that one of the cups is empty\(^{41}\). Over time, most monkeys learn to

\[ \text{Unlike the monkey study, the Mody and Carey study was set up as a competitive game. The confederate was “looking” for the cup containing a sticker, and “happened” to look in the wrong cup first.} \]
reach for the cup with the grape in it without checking to see if it is empty, or checking the other cup. Those baboons that learn to do this in this initial task (some do not, even after several weeks) move on to the next phase, where they are familiarized with a four-cup version of the same task. They are shown four cups, A, B, C, and D. Two grapes are placed in the cups, in plain view of the monkey: one in A or B, and one in C or D. A and B are placed closely together, and slightly away from C and D, which are also placed closely together. After this phase, the baboons move on to the final, test phase. Here, the four-cup task is repeated, with screens occluding the grape placement from view, so that the monkeys do not see which cups contain grape. The authors found that all of the monkeys who succeeded on the two-cup task succeeded on this four-cup task. Moreover, their learning curve for the task was shallow: The baboons acted in a way consistent with disjunctive syllogism from the first trial, and improved only marginally over time.

As Mody and Carey point out in the introduction to their paper⁴², it is important that the test condition involves four cups, rather than two. Success at the initial, two-cup phase might at first appear to suggest a form of disjunctive syllogistic reasoning. By looking in B after learning that A is empty, the monkey could be demonstrating that it has acquired the belief that the food is not in A; that it has combined that belief with the belief that the food is in A or it is in B; and that it has thereby concluded that the food is in B. But such behavior can in fact be easily explained without appeal to logical inference. The monkeys could be employing a simpler “maybe A, maybe B” strategy. In such a strategy, the monkey begins by thinking: MAYBE THERE IS A GRAPE IN A. MAYBE THERE IS A GRAPE IN B. No disjunction is represented initially, and no negation is applied upon learning that the grape is not in A. The monkey simply ceases to think MAYBE THERE IS A GRAPE IN A, while continuing to think MAYBE THERE IS A GRAPE IN B. It goes in for B, naturally: B is the only

⁴² This paper relied on 2.5 and 3 year old human participants, but the general lesson to be drawn is the same.
grape-like cup the monkey is thinking about anymore. Clearly, these two strategies lead to the same behavior in the two-cup task. However, they ought to lead to different behaviors in the four-cup task. A subject monkey in the four-cup task has evidence that two grapes are distributed across A, B, C, and D. It also has evidence that there are no grapes in A. If monkeys in this task rely on a non-logical strategy similar to the one canvassed for the two cups task, they should think something of the form: *maybe there is a grape in A; maybe there is a grape in B; maybe a grape is in C; maybe a grape is in D.* On such a strategy, the fact that A is empty tells the monkey just as much about C or D as it does about B. The updated thought, upon seeing A empty, should be (abbreviating): *maybe B; maybe C; maybe D.* The animal should be just as likely to go for cup B as for cups C or D (33%). By contrast, if the monkey was employing a logical strategy, representing ‘A or B’ separately from ‘C or D’, one would expect a different pattern of behavior. Namely, the monkey should be much more likely to go in for B than C or D. That’s because, by disjunctive syllogism, that there is a grape in A or B and that there are no grapes in A together entail that there is a grape in B. A rational agent would take the likelihood of grapes in B to be nearly certain. Cups C and D, by contrast, each present the monkey with a 50% chance of grapes.

The monkeys go for B in the four-cup task more than 33% of the time—in fact, around 60% of the time. Griffin, the parrot, was at ceiling for this task. This suggests that many baboons, as well as the parrot, were not employing a “maybe A, maybe B, maybe C, maybe D” strategy. And it supports the hypothesis that some baboons are relying on negation and disjunction in selecting action.43
This conclusion is of course very appealing to proponents of language of thought in non-human animals. Some olive baboons behave in a way that suggests their decision-making processes deploy both negation and disjunction. If their decision-making processes do that, then so do whatever aspects of the baboon mind that contribute to those processes. As argued earlier, representation of logical connectives is evidence for combinatoriality in thought. Connectives are *compositors* of complex linguistic representations. Without composition, there is no use for them. Some aspect of the baboons’ mind then, is structured in a compositional way. However, as noted earlier, even if this is the right interpretation of the data, it is not evidence for sub-sentential combinatoriality in baboon thought. That is because we do not know, on the basis of this experiment anyway, what the right semantics is for the baboons’ thought that there is grape in B.

Before getting too carried away with implications of this interpretation of the results, I describe an alternative interpretation. A different method the baboons might be employing does not rely on disjunction or negation, but is rational by the lights of probability. On that view, what guides the baboons’ behavior are not logical inferences over linguistic representations, but probabilistic inferences over non-linguistic representations (an example of this as an alternative to disjunctive syllogism is described in Rescorla 2009b). Cognitive maps could represent possible actions or outcomes (e.g, look in A, look in B, look in C, look in D). The animal would then assign a value to each map based on its expected reward. With new evidence, the probabilities could be updated, allowing the animal to select an action. For example, in the first phase of the task, a baboon might associate a probability of grape-outcome with each map, 0.5 for each. This outcome would be revised in light of evidence that A is empty, so that the probability of attaining grape for option B is near 1, while that for C and D remain at 0.5. It is important that the baboon have some way of knowing that B should be updated differently from C and D. This can be achieved if the baboon
associates probabilities with groups of cups. If the baboon ascribes a probability of 1 for attaining grape to looking in the A-B cup combination\textsuperscript{44}, then it need only “subtract” 0 from that 1 to conclude that B certainly contains grape. Alternatively, it could simply ignore A because it is empty, but still ascribe a probability of 1 to looking in the A-B combination. In that case, it would be transferring the certainty about grape in A-B to certainty about grape in B. This style of interpretation accounts for the data just as well as the logical interpretation does. Whether this interpretation or the logical interpretation is correct is ultimately an empirical question. But to the degree that one accepts the language of thought hypothesis in the human case, one should be strongly inclined accept it in this case, as well. That is because what is good for the goose is good for the gander: If you prefer a probabilistic explanation here, why shouldn’t you prefer it elsewhere?

One might think that much or all of human cognition is, to some extent, guided by computations that are strictly speaking non-logical. And the goal of this chapter has not been to convince anyone that they should not describe decision making processes in Bayesian or predictive coding terms. It has been to show that there is good evidence for a language of thought in non-human animals, and this evidence is comparable to any we have for human animals. Moreover, even if one does go this route, one would need to specify clearly, that the implementation of these Bayesian inferences does not rely on a probabilistic LoTs, an increasingly common tenet of Bayesian explanation generally (Goodman et al. 2015).

There is an additional paradigm which, though less definitive than the four-cup task, is highly suggestive of representations of disjunction in chimpanzees. In a study by Jan Engelmann (not yet published\textsuperscript{45}), the author presented chimpanzees with two cups, both of which were out of

\textsuperscript{44} In reality, it will surely be less than 1, in order to allow for noise and other sources of uncertainty. I use this probability as a toy example, but I assume that it would change little to ascribe 0.95, or 0.8, or really any value greater than 0.5

\textsuperscript{45} Engelmann has given me permission to discuss this result.
reach of the chimpanzee, and one of which always contained grape. The cups were attached to pieces of rope, the ends of which were within reach of the chimpanzee. The chimpanzee could pull either rope, or both. The cups were rigged so that by pulling one rope, the chimpanzee would easily acquire the directly attached cup, but the other cup would fall off the table it was standing on. If the chimpanzee exerted the extra energy required to pull both ropes, it could reliably acquire both cups. Chimpanzees were initially presented transparent cups, and the grape was placed in one of the two cups in plain view of the chimp. So it could be expected to be readily apparent which of the two cups contained grape. In this condition, the participants learned to pull the rope directly attached to the grape-containing cup to obtain the grape. In the test condition, the cups were taped over, so they were no longer transparent. In addition, the grape was placed in the cup behind a screen. So it could not be expected that the chimpanzees would have a good idea which cup the grape was in (though one clever chimpanzee climbed on top of the cage to look into the cups from above). In the training condition, the participants always pulled just one rope, usually the one attached to the grape-bearing cup. In the test condition, all of the participants pulled both ropes. Four of five of them did so from the first trial (the who did not was the same participant who climbed up to look into the cups from above!).

The tempting conclusion of the foregoing study is that chimpanzees represent possible future courses of action, based on possible states of the (current) world consistent with their evidence (and, perhaps, some information stored in memory). The chimpanzees are capable of representing, in some way, two possible state of affairs: that there is grape in A, and that there is grape in B. These alternatives are compared, in some way, leading to action. This is compatible with the idea that what the animals are doing is representing a disjunction: *Either there is grape in A, or there*
is grape in B. When frustrated by the fact that it cannot conclude with confidence which cup contains grape, the chimpanzee opts for the more difficult task of pulling both ropes.

As in the previous case, the chimpanzee may be using probability assignments instead of logical connectives to solve this problem. For example, it is possible that the animal associates which each possible state of affairs a probability of grape (or of positive reward). It assigns a 0.5 probability of grape (reward) to both A and B. In the absence of any new evidence shifting the chimpanzees’ probability assignments one way or the other, it opts to pull both ropes. Importantly, on this picture, the chimpanzee needn’t even represent the probabilities of grape in A and B as summing to 1. It could well be that A and B each are assigned a 0.5 probability of grape independently. It just so happens that since no new evidence leads the chimpanzee to adjust these assignments, A and B each stay at 0.5. The chimpanzee simply gives up deciding, and pulls both ropes. As in the previous example, it is also possible that chimpanzees are not representing states of affairs, such as the fact that there is grape in A, but actions, like pulling the rope attached to A.

There is an empirical test of the “non-summative” version of this probabilistic explanation that has yet to be run⁴⁶. On this explanation of the behavior, the chimpanzees do not attribute a probability of 1 to the pair of cups, but instead 0.5 to each. (By contrast, if they are attributing a probability of 1 to the pair A-B, then pulling both ropes does yield grape with certainty). The proposed test adds a new alternative for the chimpanzee to acquire grape with little chance of failure, but more effort. A second grape is placed behind a screen into a third, opaque cup, C, which is then placed farther away from the chimpanzee than A or B. In addition, C can only be obtained by pulling two ropes. So, there is as much evidence that there is grape in C as there is in A or B, but

⁴⁶ This idea came about in a discussion with Engelmann.
acquiring C is harder than acquiring A and B. Preference for any grape should decline as one increases the amount of exertion required to obtain it. But if a chimpanzee attributes 0.5 probability of grape to A and to B in a way that is not summative, then a grape that is a sure bet—even one that is somewhat far away—should be more appealing to it than a grape that is less than a sure bet. By contrast, if the chimpanzee attributes a probability of, or near to, 1 for the A-B pair, there is no condition under which a grape that is farther away—even a sure bet—should be preferable to one that is close at hand. If the participating chimpanzees tend to prefer C over A-B in this condition, this would suggest that the non-summative probabilistic answer does not explain the original result.

As a control, one would need to re-run the first experiment with three cups, instead of two, to rule out the possibility that the number of cups was influencing the outcome. In this condition non-food items would be placed in C. If the non-summative probabilistic answer explains the behavior in the experiment, the chimpanzees will opt for C just as much in the control condition as in the test condition; if the summative probabilistic interpretation, or the logical interpretation, explain it, then one will predict chimpanzees opt more for C in the test condition than in the control condition.

Even if it does rule out the non-summative explanation, this experiment would leave open whether the right explanation of the data is a summative probabilistic one or one that appeals to logical disjunction. But at this point, one can appeal to the same strategy as in the Ferrigno et al. study. We know that probabilistic accounts of various stripes can be appealed to in order to explain flexible behavior, even in the human case. So once we have reduced the question of whether chimpanzees have representation types corresponding to the logical connectives to the question of whether probabilistic or logical inferential description better explains reasoning in this case, we can declare victory.
So there is some evidence for logical connectives in non-human reasoning, and so for a kind of language of thought in non-human animals. If the logical explanation of these data is to be preferred over the probabilistic ones, then chimpanzees and baboons really do build simpler thoughts into more complex ones using disjunction. This in itself is a kind of combination, and so straightforward evidence for a language of thought. One crucial piece that is left up in the air by these results, as mentioned earlier, is the nature of the individual thoughts that are composed by logical connectives. As I mentioned before, it is possible that a logical connective, like disjunction, could be applied to thoughts that themselves have no internal compositional structure. For all we know, chimpanzees and baboons have grape-in-cup-B thoughts, rather than thoughts that there is a grape in cup B. Or, perhaps they have stored percepts of A and B, and somehow compose those into sentences, transforming them into discrete symbols. Another live question remaining is whether the competence with disjunction suggested by these data should turn out to be recursive. If the baboons and chimpanzees can apply disjunction to thoughts about A and B, then it is plausible that they can apply it to other thoughts about other kinds of thing. But can they further put together the thought that P or Q with the thought that R to conclude that P or Q or R? It is tempting to appeal to a competence-performance distinction here, and to say that they could do so, with the right working memory and attentional resources. But a conservative and troubling alternative for the LoT theorist is that there is some strange OR-like connective that turns out not to be recursive in the way that our OR concept is. One’s syntax might simply lack iterability of the OR-like concept.

Upshots for the Development of Disjunctive Syllogistic Reasoning in Humans
The foregoing discussion is in tension with the failure of 2.5 year-olds in a four cup task similar to the one baboons and a parrot succeed at (Mody & Carey 2016). In this section, I diffuse that tension. There is a likely confound in the original developmental study (ibid), which might explain the failure. Moreover, there is independent evidence suggesting (a) pre-verbal infants reason by disjunctive syllogism when given different tasks that are not susceptible to this confound; (b) children do not acquire sentential negation via acquisition of natural language connectives, the most plausible explanation of the failure at 2.5 years. Thus, it is plausible that there is a common explanation for disjunctive syllogistic reasoning in humans and non-human animals.

Here is the apparent tension between the developmental and comparative literatures. Mody and Carey (ibid) showed that while 5 year-olds, 3-year-olds and 2.5-year-olds all succeeded at the two-cup task, only 5 year-olds and 3-year olds succeeded at the four-cup task. Three years is roughly the age at which sentential negation language, such as ‘not’, begins to appear in language production. So this result is compatible with the idea that, for humans at least, natural language is necessary for competence with disjunctive syllogistic reasoning. For example, it might be necessary for acquisition of NOT. But what do we make of this result in light of the fact that, as per the foregoing discussion, adult baboons and a grey parrot appear to possess such concepts in the absence of natural language?

Here are two routes one can take the diffuse that tension. One is to accept that the cognitive development of those animals differs deeply from human beings. They acquire sentential NOT through alinguistic cognitive development, or possess it innately; we do so only or primarily with the aid of natural language acquisition. But clearly, such a stark asymmetry between us and baboons seems worth avoiding if plausible alternatives are available. Another route is to identify aspects of the task that 2.5-year old children might have struggled with, but not 3-year-olds. For example, there
is some evidence that very young children have a more limited working memory capacity than adults. Crucially, 2.5 year-olds have a working memory of just three items. Thus, the key change between the two- and four-cup tasks—the number of objects being tracked—likely makes the task much more difficult for them. If that is right, then at least as far as negation is concerned, baboon cognition might be much like ours. It is just more difficult to demonstrate that in 2.5 year olds, because of independent limitations on their reasoning abilities.

In response to the latter route, one might reply that children can manage the task by chunking the cups to get them within their working memory capacity. So they store A and B as A-and-B, and C and D as C-and-D, thereby representing two items, not four. But notice: If this is what they are doing, then they cannot be applying disjunctive syllogism to solve the task. The relevant disjunctive syllogism would involve a negation on the ensemble A-AND-B or the ensemble C-AND-D. And this won’t get us anywhere, since the important piece of information here is that there is no reward in A (Mody & Carey used stickers, rather than grapes). The resulting negation would be applied to A-AND-B. And this should lead the child to look mostly in cups C and D. But they don’t do this! Three-year-olds by and large look in cup B, and 2.5-year-olds are split across B, C, and D.

Moreover, there is independent evidence that children even younger that 2.5 succeed at disjunctive syllogism. This both supports the claim that they possess disjunction and negation and bolsters the working-
memory route to explaining the Mody & Carey results for 2.5 year olds. Cesana-Arlotti et al. (2018) showed that pre-verbal infants are surprised by outcomes of sequences that violate disjunctive syllogism. They used several different set-ups, to show the result was not a matter of the particular objects or movements used in any one manipulation. I will describe just one. 12 and 19 month olds were presented with two items, the tops of which looked just alike (e.g., a dinosaur toy and a flower). The items were occluded by a screen: the children could not see either of them (at this age, they are already disposed to think that objects that pass behind other objects continue to exist, even though they are no longer visible). Then, an experimenter takes a bowl and places an object in it behind the screen. When the bowl is brought back from behind the screen, only the top part of the object in the bowl—the part that looks just alike in both toys—is visible. Thus, it is reasonable to assume that either object may be in the bowl: flower or dinosaur. The screen is then removed, so the child can see the object that was behind the screen. At this point, application of disjunctive syllogism could easily lead one to conclude that whatever is in the bowl is the other of the two objects: For example, if there is a dinosaur behind the screen, one can conclude that there is a flower in the bowl. Then, screen and fully visible object are removed, leaving only the object in the bowl. The object in the bowl is lifted out of the bowl. In half the trials, the object in the bowl is what is predicted by disjunctive syllogism (e.g., a flower). In the other half, it is not: While the children cannot see what is behind the screen, the experimenters switch out one of the objects (e.g. a dinosaur) for another (e.g. a flower). The authors report that infants looked for a longer time at the revealed object the second condition. This suggests the outcome violated the expectations of the infants. The most straightforward explanation of this surprise is that they executed a disjunctive syllogism at some point after the object behind the screen was revealed, but before the object in the bowl was revealed. Since the infants have not begun producing natural language negation or disjunction yet, this
suggests that they have some competence for this inference before language acquisition. And it undermines the alternative to the working memory route in explaining the Mody & Carey result.

The claim that performance constraints inhibit application of negation in 2.5 year-olds is also indirectly supported by a clever observation of linguistic production of negation. Jesse Snedeker and Roman Feiman (poster at 2019 SPP) looked at frequency of production of natural language sentential negation words (such as ‘not’) in Russian- and Chinese-born children adopted in the U.S. These children are less far along than their U.S. peers in native English acquisition. But naturally, they are on par with them regarding cognitive development. Thus, they offer a perfect way to cleave apart the hypothesis that negation is acquired by natural language—the most appealing alternative to working memory constraints for explaining the 2.5 year old’s failure—from the hypothesis that it stems from some other feature of cognitive development. If language acquisition, rather than cognitive development, is the source of competence with linguistic sentential negation, we should expect their production of sentential negation expressions, like ‘not’, to emerge at roughly the same stage of language development as linguistically-matched peers. But this is not what happens: The Russian- and Chinese born children began producing linguistic sentential negation at the same time as their peers matched for age.

So it looks as though some other feature of their cognitive development places an early constraint on expressions of sentential negation. This might be due to lacking certain concepts (perhaps, somehow, they lack NOT but acquire it later). Or it might be due to an inability to combine certain concepts (perhaps there is a very limited ways they can use NOT at that age\footnote{This would be fairly unsurprising, given the degree to which negation induces load in adults, making understanding sentences tasks more than affirmative sentences (Wason & Johnson-Laird 1976; for a longer discussion, Mandelbaum 2013a). Cognitive load might impede combination.}). Or, perhaps
most oddly, it might be due to a non-linguistic constraint on which mental states they are able to express. This is only indirect evidence for the working memory route to interpreting the Mody & Carey result: It undermines the most plausible explanation for what would explain the 2.5 year old's failure, if not some performance constraint. Since sentential negation in thought likely does not come from sentential negation in language, it cannot be that the 2.5 year olds fail because they lack natural language negation. (And, if anything, the correlation is explained by causation in the other direction: a lack of negation in thought explains the lack of negation in language)

It looks as though the initial failure of 2.5 year olds reported by Mody & Carey reflects a performance constraint on applications of logical inference in those children, rather than evidence for a link between language acquisition and logical thought. Human development, like that of baboons and chimpanzees, leads to the mastery of at least one form of logical inference in the absence of language acquisition. The plausible and exciting upshot of all this is that there are important similarities in the nature of the competence for logical inference that we share with those animals. It bears remarking that this strengthens the overall abductive argument of this dissertation. LoTs can explain this commonality of rational capacities. Moreover, the plausible assumption that our LoT shares some syntactic features with that of other primates yields the direct prediction that we will all execute representational transformations of similar form. This includes deductive inferences. The aligning of this prediction with the observed commonality strengthens an inference to the best explanation that relies on the widespread presence of LoTs in animal minds.
What Was at Stake?

To conclude, it might be useful to remember what was at stake. I’ve presented some evidence that grey parrots and olive baboons are capable of disjunctive syllogism, and potentially chimpanzees as well. If it is true that they possess these states, then there is a deep structural continuity in the kinds of mental state that such animals can have and the kinds we have. To the degree that LoT’s really do underpin language, perception, thought, and rational inference, it is conceivable that some animals, including the ones mentioned above, possess a kind of *competence* for these capacities, even if many of them never in fact realize this competence. Finally, and perhaps most interestingly, the studies I’ve described here offer an exciting proof of concept for diagnostic tools for the representation of logical connectives in non-human animals—and so for a kind of LoT in those animals.

This evidence also goes part of the way in responding to some of the negative arguments considered in Chapter Three. The conjecture that the selective advantage of language is such that any creature with a language of thought ought to acquire natural language is all the more damaged by this independent evidence for non-human LoT’s. That’s because that argument was only suggestive of the implausibility of the LoTH for non-linguistic animals in virtue of the lack of independent evidence of for LoT’s in those animals. This opens up the possibility that many other animals who do not possess natural language nonetheless have mental states that are structured in a language-like way.
Similarly, Liz Camp’s rejection of Cheney & Seyfarth’s argument for a baboon LoT is weakened by this evidence. It is still possible that baboons represent dominance hierarchies (and potentially other features of their environments) in a map-like way. But we now have some independent evidence for combinatorial structure in baboon thought. This supplemental evidence weighs in favor of the Cheney-Seyfarth interpretation of baboon social behavior. It strengthens their original inference to the best explanation against Camp’s alternative proposal.

The LoTH for perception is weakened by the possibility that the representations that feature in perceptual computations are picture-like, rather than language-like. But it is strengthened, however weakly, by the discovery that language-like representations do appear elsewhere in psychological explanation for at least some non-human animals. At the very least, one cannot reject the LoTH for chimpanzee and baboon perception on parsimony grounds.

The argument from concept learning in non-human animals is weakened by the possibility that non-human animals’ concepts lack the compositional structure of a LoT. If they do have such a structure, we can conclude that any animals who can learn concepts have a compositional structure to their thoughts. But we still cannot conclude that their thoughts can be combined into complex thoughts, to serve in inferences—a key feature of the LoTH. The discovery that chimpanzees and baboons possess mental representations corresponding to disjunction and (in the baboon case) negation offsets at least the second of these two shortcomings. So, if chimpanzees, baboons, and grey parrots (and any other animals that pass these tests), have concepts that are compositional, then concept learning is evidence for LoTs that are more familiar to us from discussions of human cognition—ones with both sub-sentential and sentential combinatoriality.
I stress, also, some things that these results do not show. Nothing about these results, of course, bears on connectives other than disjunction or negation: conjunction, material conditional, or biconditional. And so it is possible that practical deliberation in chimpanzees and baboons includes disjunctive syllogism, but lacks, for example, modus ponens, modus tollens, or even simplification (P and Q, therefore P). Of course, the other connectives may be derived, formally, from disjunction and negation. For example, these animals might be capable of inferences formally equivalent to modus ponens using only negation and disjunction. But there need not be any unique LoT symbol that represents the other connectives. And anyway, it is an empirical question whether any animal with negation and disjunction does compose those symbols with sentences to perform those inferences. It is also unclear whether the LoTH of any animals includes quantifiers. Finally, even if the supposition that some animals have subsentential structure to their thoughts that includes predicates and arguments, it is unclear what the adicity of those predicates looks like; or whether there is some kind of syntactic upper limit for composition of those predicates.
Chapter Six: Conclusion

What are my dog’s thoughts like, and how might I go about figuring that out? The suggestion I proposed, for my dog and for all of the other species one might be curious about, was to take seriously the possibility that some creatures possess a LoT. This was the challenge that opened this dissertation. I have described some different ways that minds might possess a LoT. And I have argued on empirical grounds that six species, from far apart clades of the phylogenetic tree, look very much as if they possess them.

Despite the obvious appeal that the LoT has as in explaining the mental capacities of non-human animals, the familiar arguments in favor of the LoTH do not seem to support this strategy. This leaves us with only a meager abductive strategy: Garner a large body of evidence, from all sorts of species, and show that a LoT is needed to explain it.

Accordingly, I have laid out evidence for a range of non-human LoTs. On one end of the spectrum, there is the rudimentary mind of the fruit fly, which, as far as we can tell, lacks combinatoriality. On the other, Chimpanzees, baboons, and a grey parrot appear to possess a capacity for rational thought human exceptionalism once took as a point of pride. This suggests they possess a LoT with (at least) sentential connectives of some kind and, potentially, a gamut of other psychological competences. Somewhere along the line fall the paper wasps, the honeybees, and the bumblebees, who seem, at the very least, to be able to combine mental representations. We still do not know how much of their environment these animals really represent, or indeed how complex
the syntax of their minds might be. But the mere discovery of syntax in such small nervous systems is, I take it, an accomplishment of the LoT strategy I set out to apply.

I take this dissertation to be part of a larger project of studying non-human mentality. Chapter Two laid out some syntactic and semantic axes along which LoTs might differ. This helped us, not just to get clear on that posit, but to exercise caution in interpreting the relevant evidence for non-human LoTs. However, it did not exhaust those axes. There are likely many kinds of combinatorial grammars that were not considered there. Accordingly, there may be a large number of grammars that a combinatorial mind might implement.

The empirical arguments in Chapters Four and Five are strongest when taken together: A single theoretical tool can explain significant aspects of both wasp mentality and chimpanzee mentality (this is to say nothing of homo sapiens). This strengthens the case for the explanatory role of this psychological kind. But it also underscores the small number of species, and pieces of evidence, discussed in this dissertation. There is much to be gained from considering a larger number of species, and in more depth. (This will depend, in part, on more evidence emerging). Such evidence could, of course, either strengthen or weaken the central thesis of this dissertation. But it is also sure to give us a better picture of representational combination as a psychological trait: how different kinds might have emerged over evolutionary history, how similar variations of it might be across taxa, whether it is a convergently evolved trait, what the most primitive forms of it are, and what if anything makes the human variety special.
References


Block, N. (not yet published) “Perception is Non-Propositional, Non-Conceptual and Iconic” Chapter 3 of unpublished manuscript


Engelman, J., Völter, C., O'Madagain, C., Proft, M., Rakoczy, H., and Herrmann, E. Title information is not available, and the paper has not yet been published


Ferrigno, J., Huang, Y., and Cantlon, J. Title information is not available, and the paper has not yet been published.


Fuss, T., & Schluessel, V. (2017). The Ebbinghaus illusion in the gray bamboo shark (Chiloscyllium griseum) in comparison to the teleost damselfish (Chromis chromis). *Zoology, 123*, 16-29. doi:10.1016/j.zool.2017.05.006


Pack, C., & Theobald, J. (2018). Fruit flies are multistable geniuses. PLOS Biology, 16(2). doi:10.1371/journal.pbio.2005429


