9-2017

Syntax and Semantics of Perceptual Representation

James K. Quilty-Dunn
The Graduate Center, City University of New York

SYNTAX AND SEMANTICS OF

PERCEPTUAL REPRESENTATION

by

JAMES K. QUILTY-DUNN

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

2017

Syntax and Semantics of Perceptual Representation

by James K. Quilty-Dunn

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

_____   _____

Date   Eric Mandelbaum

Chair of Examining Committee

_____   _____

Date   Iakovos Vasiliou

Executive Officer

Supervisory Committee:

Jesse Prinz

Eric Mandelbaum

David Papineau

Tatiana Aloi Emmanouil

Ned Block

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Syntax and Semantics of Perceptual Representation

by

James K. Quilty-Dunn

Advisors: Jesse Prinz & Eric Mandelbaum

The notion of representation is fundamental to cognitive science. Representations figure in psychological explanations of low-level perceptual processes like shape perception as well as the more "central" reasoning processes studied by cognitive and social psychologists. A fault line for theories of the architecture of the mind concerns whether perceptual processes deliver representations couched in special modality-specific formats, or whether perception deploys amodal conceptual states also used in cognition. This dissertation is a defense of *perceptual pluralism*, according to which perception delivers both kinds of representations. I develop and motivate perceptual pluralism with particular attention to its consequences for the border between perception and cognition.

Converging empirical evidence in several areas (e.g., mental imagery, high-capacity sensory memory, the structure of visual cortex) suggests that some perceptual representations are *iconic*, or image-like. Iconic representations are syntactically and semantically distinct from *discursive* representations such as sentences and propositional thoughts. For example, unlike

propositional thoughts, icons cannot explicitly predicate properties or figure in classical computational processes. I develop a compositional syntax of icons called the *coordination model*, according to which each part of an icon determines values along multiple feature dimensions (such as color and spatial location) simultaneously. The coordination model departs from rival feature-placing models of perception and imagery. I show how the model can be used to develop a semantics of icons, and how it can help explicate the notion of nonconceptual content.

Not all perceptual representations are iconic, however. Object perception—the ability to perceive and track objects by means of the senses—is a rich perceptual capacity present in both infants and adults, and has been the focus of intense study in perceptual and developmental psychology. I argue on empirical grounds that perceptual object representations (PORs) have a discursive, non-iconic format that involves distinct representations for separate individuals and their properties. Furthermore, PORs are amodal, represent high-level properties, compose in concept-like ways, are mapped to lexical representations, are propositionally structured, and are apt to function in inferences. I argue on these grounds that PORs, like beliefs and other cognitive states, are composed of concepts.

Perceptual pluralism has consequences for theories about the distinction between perception and cognition. Theorists like Tyler Burge, Susan Carey, and Ned Block argue that this distinction lies in part in the fact that perceptual states are iconic and nonconceptual. The fact that PORs are conceptual and not iconic, however, is incompatible with this approach to the perception–cognition border. One option, then, is simply to deny that there is any such border. Another option is to appeal to architectural factors, such as modularity. Defenders of modularity

(such as Firestone & Scholl) have often argued that perceptual modules are immune to the influence of cognition over time. However, PORs contain concepts like PIANO and FISH that may not be innate, suggesting that learned concepts can seep into the object-perception system.

I sketch an architectural approach to the perception–cognition border, according to which perceptual processes (including object perception) are *synchronically encapsulated*; that is, they operate on a fixed set of information at one time, which does not include beliefs, desires, or other information stored in cognition. This set of information can change over time, however, to integrate new concepts through learning. The resulting view incorporates key insights of modularity theory by holding perception to be synchronically encapsulated and dependent on sensory stimulation, while also incorporating anti-modularist insights by allowing the outputs of perception to be conceptual and for perception to be shaped by new concepts over time.

Dedicated to Declan Michael Dunn and the memory of Michael Edward Dunn

ACKNOWLEDGEMENTS

Ross Colebrook has been a regular source of philosophical insight and clarity. I'm grateful to Ross for being a good friend and for being simultaneously one of the funniest and the most morally serious people I know. My time in grad school would have been much more difficult and less interesting without him.

Thanks to E.J. Green for being a wellspring of knowledge and ideas about both the experimental research and philosophical issues discussed below. Being able to talk to E.J. about perceptual psychology was one of my secret weapons in writing this dissertation. E.J.'s love for the music of Billy Joel has also served as a much-needed reminder that exceptional philosophical acumen does not necessarily correlate with good judgment in other domains.

I am very grateful to the members of my dissertation committee: David Papineau, Tatiana Aloi Emmanouil, and Ned Block. They have provided excellent feedback and support. All three are models of, among other things, how to participate in a broader research community in a productive and enjoyable fashion. Ned's work plays a central role in this dissertation and conversations with him have had a major impact on the material below, for which I am extremely grateful. Participating in lab meetings under the guidance of Tatiana as well as Steve Young, and learning from both of them, has been one of the most fruitful experiences I've been fortunate enough to have in grad school.

There are three philosophers without whom this dissertation wouldn't have been completed.

I have been unfathomably lucky to have the two advisers I have had. Jesse Prinz has been an invaluable source of knowledge, support, criticism, and clarity. I am convinced that Jesse has somehow figured out the secret to total recall before the technofuturists; I know of no other earthly explanation for the wealth of experimental, philosophical, and cultural knowledge he carries in his mind. His ability to use that knowledge in such creative, subtle, and insightful philosophical ways, including in advising this dissertation, confounds me further. I'm extremely grateful to Jesse for his guidance and the example he provides of how to be a philosopher and teacher.

Eric Mandelbaum is a brilliant and thoughtful philosopher, a wise and patient teacher and adviser, and a very good friend. Like Jesse, Eric possesses a brain that houses enough information to make Google envious, and deploys it in ways that are both clever and deep. I cannot overstate how much I have learned from Eric and how grateful I am to have been able to work with him. In addition to his philosophical guidance, he has been a constant source of sanity, rationality, and compassion, for which I am thankful.

Zoe Jenkin has provided an endless supply of critical philosophical insight as well as loving support. This dissertation would look nothing like it does without the former and would not exist at all without the latter. Her brilliance and warmth have sustained me through the morass of grad school. I am immensely lucky to know her.

Lastly, I owe thanks to my family (including both the Quiltys and the Dunns) for their support. My mother Beverly and my stepfather Colin have been incredibly supportive and caring, for which I am grateful. Arguing for hours on end with my perceptive and fiercely intelligent

brother Liam during my formative years is probably somehow responsible for my entering

philosophy in the first place and I am thankful to him for that.

TABLE OF CONTENTS

xiv

LIST OF FIGURES

"The log of knowledge or information contained in the brain of the average citizen is enormous. But the form in which it resides is largely unknown."

    —Cormac McCarthy

"There are two kinds of visual memory: one when you skillfully recreate an image in the laboratory of your mind, with your eyes open (and then I see Annabel in such general terms as: "honey-colored skin," "thin arms," "brown bobbed hair," "long lashes," "big bright mouth"); and the other when you instantly evoke, with shut eyes, on the dark innerside of your eyelids, the objective, absolutely optical replica of a beloved face, a little ghost in natural colors[.]"

    —Vladimir Nabokov

"Philosophy of perception must become largely a branch of philosophy of science. Since it must, it will."

    —Tyler Burge

INTRODUCTION

It is evident that thinking about tigers and seeing one in the flesh are not the same sort of mental process. It is less evident how precisely they differ. This dissertation is an attempt to articulate aspects of this difference in the vocabulary of cognitive science. Understanding the border between perception and cognition is a first step toward understanding the architecture of the mind more generally. While a reader will not come away from this document with a comprehensive understanding of the perception–cognition border, they will hopefully come away with some more clearly formulated opinions on the subject. Ideally those opinions will align with the ones put forth below, but the primary goal is to get clearer on the questions rather than to provide definitive and uncontroversial answers.

The core thesis is *perceptual pluralism*, according to which perceptual systems deploy a multiplicity of representational forms including both proprietary (iconic) representations and non-proprietary (discursive, conceptual) representations. The heart of the view is articulated in Chapter 2; Chapter 3 extends the core argument and Chapters 1 and 4 clarify some foundational issues in the neighborhood. Nonetheless, the views articulated about the vehicle-content distinction and informational encapsulation in Chapter 1 as well as the coordination model of iconic representation defended in Chapter 4 float free of the truth of perceptual pluralism. Some approaches to the perception–cognition border are ruled out by perceptual pluralism while others are highlighted as avenues for further investigation. The Conclusion points toward future research on this topic.

This dissertation is technically a work of philosophy. The aim is to understand the human mind, an aim that belongs to no particular academic discipline. Much of the discussion is relatively *a priori*, e.g., about what does or doesn't count as a concept or as an iconic representation. Many of the citations in the bibliography are psychology experiments. There are no uniquely philosophical methodologies employed: no conceptual analyses, no delineation of synthetic *a priori* truths, no phenomenological reductions, no careful attention to ordinary language. Instead my goal is to make evidence-based claims about the representations delivered by human perceptual systems and explore how the truth of those claims bears on the most basic structural aspects of the mind.

Because this project involves making large-scale theoretical hypotheses that synthesize a wide variety of evidence, and because these hypotheses bear directly on some central debates in the history of philosophy, it seems appropriate to call this dissertation a work of philosophy. But as Jerry Fodor wrote about whether a certain type of theory counts as linguistics, "The question what it is for a linguistic theory to be true is an *interesting* question[.]…The question what it is for a true theory to be linguistic is a *boring* question" (1981b, 197). The question of whether a theory is part of philosophy or part of cognitive science (if those are mutually exclusive) is a boring question. The questions of how perceptual representations are structured, how perception differs from cognition, and what distinguishes the various representational formats and computational systems of the human mind, however, are interesting questions.

MENTAL ARCHITECTURE:

IN SEARCH OF A BLUEPRINT

Chapter 1

"You need a blueprint, a recipe, an instruction manual, a program. This goes for the mind as well as any other contraption."

> *—Fred Dretske[1]*

§1. Mental Architecture as the Fundamental Object of Cognitive Science

Cognitive science, whatever it is, aims at understanding the most basic aspects of the mind. And the mind, whatever it is, is probably not a single undifferentiated thing but instead likely consists of distinct interacting parts. A primary goal of cognitive science is thus what Hume called a "mental geography, or delineation of the distinct parts and powers of the mind," which may enable us to "discover, at least in some degree, the secret springs and principles, by which the human mind is actuated in its operations" (Hume 1748, 1.1, SBN 13–14).

A theory of mental architecture (or "geography") should specify what the simplest parts of the mind are, how they operate, and how they interact with one another. At the heart of mainstream contemporary approaches to mental architecture lies the notion of representational mental states, or *mental representations*. Mental representations are states that have content—that

---

[1] Dretske 2000, 208.

are directed at things beyond themselves—and that may be manipulated in ways that instantiate mental processes such as thinking, acting, remembering, imagining, and perceiving, as well as more low-level processes such as stereoscopic depth computation or parsing the syntactic structure of a sentence. A standard way of understanding mental representations in philosophy and cognitive science is to construe them as functionally individuated mental particulars and to construe mental processes as instantiated by computational operations on those mental particulars (e.g., Fodor 1975; 1987; 2000; Chomsky 1980; Marr 1982; Fodor and Pylyshyn 1981, 1988; Pylyshyn 1984; Pinker 1997; Rock 1997; Palmer 1999; Prinz 2002; Carey 2009; Frisby and Stone 2010).

Below I will often refer to mental representations, mental processes, and mental systems. A representation is a contentful state. Mental processes are operations performed on representations. Processes tend to transform representations in a fashion that realizes a particular rule-governed input-output function; that is, mental processes tend to be computational. Perhaps not all mental processes are genuinely computational but I will assume that the processes under discussion here are computational unless the evidence demands otherwise. A mental system is harder to define but I will understand a system as an organization of processes that maps onto an explanatorily useful division at the level of mental architecture.

One way of understanding divisions between systems is that a particular system consists of a process or group of processes together with their shared resources and information that those processes use to compute their outputs, whether that information be stored in the form of explicit symbols or be simply built into the processes themselves (i.e., in the form of rules). The object-perception system for example contains information about how objects are structured and move

(Pylyshyn 2007). That information is used in individuation and tracking processes, which (respectively) construct representations of objects and update them as those objects move. This information is shared across processes and representations used in one process might be used in another as well (e.g., the output of the individuation process is the object representation used for tracking). On this notion of systems object individuation and object tracking are part of the same object-perception system.

There are doubtless other joints at which one could carve the mind but the distinction between systems, processes, and representations will prove useful.

Some theorists adopt anti-realist or instrumentalist positions about mental representations generally (e.g., Churchland 1981; Dennett 1987; 1991) or about perceptual representation in particular (e.g., Gibson 1979; Campbell 2002; Noe 2004; Brewer 2006; Martin 2006). Others accept that mental states represent but deny that they can be given functionalist identity conditions or that they figure in computational processes (e.g., Searle 1983; 1992). Still others accept that there are functionally individuated particulars that figure in computational processes but deny that they have contents (e.g., Stich 1983, Chomsky 2000). I assume on the contrary that mental representations are real, have contents, and figure in computational processes. I will not attempt to justify these assumptions except by demonstrating their usefulness in theorizing about the mind. Such a demonstration is arguably as good a justification as one could hope for in the realm of theoretical posits.

Some realists about mental representation adopt a *globalist* attitude toward mental processes and systems. Globalists hold that there are no real borders to be drawn between mental

systems and/or that a small number of mental processes are psychologically ubiquitous and explain all or most mental phenomena. A classic example of this is unreconstructed associationism, according to which all mental processes are reducible to associative connections between representations (though interestingly Hume seems to endorse robust distinctions between mental systems despite his associationism).

Another family of globalist theories currently *en vogue* is the motley crew of Bayesian and predictive coding approaches to the mind, according to which all mental processes are probabilistic computations. On the strongest versions of this view architectural borders between mental systems should be replaced by a free flow of hierarchies of information processing replete with top-down and bottom-up influences (e.g., Clark 2013; 2016; Hohwy 2013; Lupyan 2015).

Another stripe of theorist accepts architectural borders between systems but denies a plurality of representational kinds. These *anti-pluralists* argue that all mental representations are couched in the same "common code" and that any apparent differences are to be explained by appeal to architectural borders between systems or differences in the kinds of processes that operate on representations in this common code. One prominent anti-pluralist is Zenon Pylyshyn (1984; 2003). Pylyshyn argues that all mental representation takes place in an amodal language of thought and that processes like mental imagery and perception differ from ordinary thought only in respect of the computations they perform and the systems in which they take place.

Another anti-pluralist is Jesse Prinz (2002). Prinz endorses a form of concept empiricism according to which concepts are stored copies of perceptual states and perception differs from thought only in its being stimulus bound (though unlike Pylyshyn Prinz posits distinct formats for

different perceptual modalities). Prinz (2006; unpublished) also seems to endorse a mitigated form of globalism that denies architectural borders but allows for a thin distinction between processes by appeal to functional decomposition.

I'll argue later for a view that is both *anti-globalist*—in that it recognizes borders between systems and a diversity of mental processes—and *pluralist*—in that it recognizes multiple types of representations. For now, however, I simply flag that globalist and anti-pluralist views are not mandatory and no such framework will be assumed here.

With respect to the border between perception and cognition most of the debates I will be engaged in are in-house disputes between anti-globalists (which is of necessity given that globalists typically reject the border in question). Among anti-globalists, there are two general strategies for distinguishing perceptual systems from central cognition: the representational strategy and the architectural strategy. The representational strategy seeks to show that perceptual representations are distinct from the representations used in thought (Dretske 1981; Tye 1995; Smith 2002; Carey 2009; Burge 2010a; 2010b; 2014a; 2014b; Block 2014; unpublished). The architectural strategy on the other hand seeks to show that perceptual processes(/systems) are distinct from cognitive processes(/systems) (Fodor 1983; Pylyshyn 1984; 1999; Firestone & Scholl 2016).

The representational strategy commits to pluralism about mental representation generally but it commits to an anti-pluralism about *perceptual* representation—no perceptual representations can, on this strategy, be type identical to those used in thought. The architectural strategy makes no such commitments. It allows for domain-general anti-pluralism (Pylyshyn 1984; 1999) but it equally allows for radical pluralism. Unlike the representational strategy, the

architectural strategy allows for *perceptual pluralism*, or pluralism about the representational forms used in perception.

This dissertation is a defense of perceptual pluralism. My central goal is to provide a detailed and well-motivated characterization of some of the outputs of perceptual processing particularly insofar as the nature of the outputs of perception informs more general questions about mental architecture. I will not defend the architectural strategy directly, though later in this chapter I will sketch a promising version of it and point out ways in which some cross-talk from cognition to perception is compatible with architectural borders. However, the representational strategy is not compatible with the form of perceptual pluralism I will defend. If the perception–cognition border is to be defended it should therefore be defended through the architectural strategy or perhaps some other strategy entirely.

For the rest of this chapter I'll explore some basic questions about mental architecture as they relate to the border between perception and cognition. These include the vehicle-content distinction (Section 2), translation between perception and cognition (Section 3), and the nature of informational encapsulation and its potentially fundamental role in mental architecture (Sections 4, 5, and 6).

§2. Representations as Vehicles for Content

2.1—Representations and dispositions. In the previous section, systems were framed partly in terms of processes and processes were framed partly in terms of representations. This framing presupposes that mental representations are in some sense the most primitive elements of mental

architecture. The view that mental representations are metaphysically more basic than mental processes has a checkered past.

Jerry Fodor writes that the "characteristic doctrine of twentieth century philosophy of mind/language" (2004, 29) is that having concepts requires having epistemic dispositions. Many analytic philosophers have held that one could not count as having concepts if one could not perform inferences with those concepts (e.g., Sellars 1956; Quine 1960; Davidson 1967; Dummett 1973; Evans 1982; Peacocke 1992; Brandom 1994; Burge 2010a—and possibly Wittgenstein). Others, most prominently Fodor, have insisted instead that concept possession is simply a matter of being able to deploy a certain kind of mental representation and does not require inferential capacities.

I don't intend to step into this debate. I do however want to target a metaphysical assumption that is often intertwined with the "characteristic doctrine," viz., the rejection of the metaphysical priority of representations. One could endorse the claim that a representation counts as conceptual only if it plays a role in inferences but also insist that representations are metaphysically prior to inferential processes. Endorsing the necessity of inferential capacities for concept possession does not entail that mental representations are *mere dispositions of the thinker* to engage in inferences. This point applies to mental processes generally: stipulating that mental representations always figure in processes does not entail that representations are mere dispositions for processes to occur.

The thesis that mental representations are mere dispositions of the creatures that have them raises serious problems. The fact that mental representations figure in computations allows for

satisfying causal-mechanical explanations.[2] Causal-mechanical explanations of behavior in perceptual psychology involve the following sort of causal chain: some process takes transduced stimulation as input and delivers a mental representation as output; this mental representation enters into computational processes the end result of which is an action (such as the pressing of a button or a verbal report). It is an essential feature of these explanations that token mental representations are concrete particulars that cause other token representations and behavior.

Saying that a mental representation is a mere set of dispositions—such as dispositions of the organism to behave in some way or dispositions to have thoughts on the basis of perception—precludes this kind of explanation (Strevens 2008). A set of dispositions is not a concrete particular that can be an input to a causal computational process. Reducing mental representations to dispositions to engage in inferences or other mental processes thus precludes a causal-mechanical explanation of those processes. One could retain a causal-mechanical explanation of processes on a dispositionalist view by holding that the categorical bases of dispositions have causal powers. However, on the view on offer those categorical bases could not be mental representations (since mental representations are *ex hypothesi* mere dispositions). Instead the categorical bases must be neural, in which case there is no psychological-level causal-mechanical explanation (see Quilty-Dunn & Mandelbaum unpublished a). For there to be psychological-level causal-mechanical

---

[2] I don't intend to take on a full-bore commitment to causal-mechanical explanation as the essence of scientific explanation (cf. Salmon 1984; Strevens 2008). All I claim here is that satisfying explanations in the special sciences typically invoke causal mechanisms of some sort and specifically that computational-representational explanations in psychology should be understood in causal-mechanical terms.

explanations, mental representations can't simply be dispositions for mental processes to occur; they must be particulars that figure causally in mental processes.[3]

A related worry that pushes toward the same point is that computational processes are characterized in terms of the representations they operate on. If that's true, then the view that representations are mere dispositions for processes to occur cannot be formulated coherently without circularity (Fodor 2004).

This is not to deny that functional roles can be constitutive of types of mental representations (e.g., Lewis 1972; Fodor 1975; Dretske 1981). But however functional role figures in individuating representations, a representation cannot be nothing more than the disposition for some process to occur. I will discuss other ways of individuating representations below. That discussion, however, will presuppose the distinction between vehicles and contents, to which I turn now.

2.2. The vehicle–content distinction. Token representations should be understood as concrete mental particulars that have content and figure in mental processes. Moreover, just as representations don't reduce to the processes they figure in, they also don't reduce to their contents. It's obvious that representations in the ordinary sense (e.g., pictures, maps, sentences) are not identical to their contents. Consider the following sentences:

(1) Snow is white.

(2) Schnee ist weiss.

---

[3] This argument really targets "superficial" versions of dispositionalism (Schwitzgebel 2013); there may be sophisticated versions of dispositionalism that not only avoid this objection but actually comport with one of the views sketched below.

(1) and (2) ostensibly have the same content, viz., that snow is white. Nevertheless, (1) and (2) are not two tokens of the same sentence. Moreover, a photograph of snow might represent it as white and thus seemingly carry the same content while still being identical to neither (1) nor (2). A representation therefore does not reduce to its content. These quotidian examples establish a distinction between contents and vehicles. A representation is a vehicle with content; or as is sometimes said, a symbol with an interpretation (Pylyshyn 1984). (1) and (2) share content but have distinct vehicles for that content.

The vehicle–content distinction applies to mental representations as well. For example, one can represent a red apple in thought or in visual imagery. Both these mental representations arguably share the content <red apple> but represent that content via distinct vehicles. I'll discuss the experimental literature on mental imagery in some detail in Chapter 2. But for now, an intuitive example may suffice. Consider the following question: what color is a bumblebee's head? You might be an expert on bumblebees and have the representation BUMBLEBEES HAVE BLACK HEADS stored in central cognition.[4] But if not, you might answer this question by forming a mental image of a bumblebee and inspecting its head, realizing thereupon that the head is black (Kosslyn 1980, 113). In both cases the content that bumblebees have black heads is stored but the vehicles for that content seem to be different, similarly to how a photograph and a sentence can carry the same content via distinct vehicles. Exactly how these representational formats differ will be a major topic of the rest of this dissertation but at present we can simply note that the difference seems to be at

---

[4] I will use small caps to denote concepts and conceptually structured representations.

least partly in the vehicles of the representations—an intuition that will be given much more substantiation in later chapters.

Another family of examples are known as "Frege cases." A person might have a mental representation of the original host of the reality show The Apprentice and a mental representation of the 45th president of the United States and fail to realize that those are, bizarrely enough, the same man. Since these mental representations have distinct computational roles in that person's mind they plausibly constitute different vehicles for the same referential content (Fodor 2008).

In both sorts of cases one might find it plausible that there is also a difference at the level of content. Frege cases in particular have motivated Fregeanism about content, according to which differences in cognitive significance like those just mentioned involve different Fregean *senses*, which are ways of presenting referents. For Fregeans these modes of presentation constitute the semantic values of linguistic and mental representations. I don't intend to assume here that Fregeanism is false (my thoughts on this topic change with the winds). Even if Fregeanism is true the differences in content in these examples also need to involve a difference in vehicles for those contents.

An extreme Fregean might balk at this claim and instead hold that all structure invoked in representational psychology is in terms of content and deny that vehicles have any explanatorily significant structural features. There are two ways of articulating this extreme Fregean view.

The first holds that all vehicles of content are unstructured atoms and that the only structure in the mind is present in Fregean senses. That is, the content of the belief that *a* is F might be a structured Fregean sense consisting of the senses of *a* and of F as well as whatever unifies them

into a whole propositional content (though specifying this last part is no mean feat—Jespersen 2012). But the vehicle is on this view an unstructured entity that is marked only by its carrying this Fregean content. It is not always clear when this view is being endorsed, but it may be at least implicitly present in the work of Gareth Evans (1982) and Christopher Peacocke (1992).[5]

The second version of extreme Fregeanism simply denies the vehicle-content distinction, and asserts instead that there are only types and tokens of structured Fregean content. On this view, there are no non-semantically individuated properties of mental representations (aside from representationally irrelevant properties of tokens, e.g., blind neural properties). It is again not clear when this view is endorsed, but on one reading it is present in Burge (2010a)—see the following two chapters for more discussion on the vehicle–content distinction in Burge's work.

Neither of these views should be accepted. It is plausible that content is at least typically *wide*, i.e., it depends on causal relations to the environment. Computation requires the computationally efficacious properties of a mental representation to be accessed by the relevant computational process. Those properties must therefore be local, i.e., internal to the computational system. If contents are typically wide then computational systems typically do not have access to contents; thus computational systems must be able to access non-semantic properties of mental representations. Moreover, it is hardly a coincidence that computation outside the mind is cast in formal, syntactic terms. The syntactic (i.e., vehicular) properties of mental representations are both local and non-semantically individuated and are therefore poised to be the computationally

---

[5] Though Peacocke (2004) does later seem to endorse the language of thought hypothesis.

efficacious properties of mental representations (for more argument see Chapter 4, ICONIC REPRESENTATION).

This point is meant merely to establish the need for a vehicle–content distinction in computational psychology. But it does not establish the need for *structural* features of vehicles, i.e., syntactic structures. The debate about whether mental representations have syntactic structures has been long and hard-fought (e.g., Fodor 1975; 1987; Chomsky 1980; Rumelhart & McClelland 1986; Fodor & Pylyshyn 1988; Smolensky 1990; Prinz 2002; Carey 2009; Burge 2010a).

The debate about syntactic structure has often been founded on systematicity, or the capacity to recombine thought contents systematically, and productivity, or the capacity to form new thoughts (Fodor & Pylyshyn 1988). Arguably the dominant view explains these phenomena by appeal to word-sized concepts that compose into more complex concepts and into propositional thoughts. The challenge from the orthodoxy is to provide a satisfying explanation of these commonplace features of thought without adverting to syntactic structure. I will not weigh in on the present status of this debate here except to note the lack of a clear and widely endorsed alternative viewpoint. One might completely reject the Fodorian language-of-thought picture and nonetheless be moved to hold that concepts compose into complex representations with syntactic structures (e.g., Prinz 2002). There are moreover many other phenomena that can be explained by syntactic structure.

One example is logical inference, which appears to utilize the propositional syntactic structures of thoughts (Fodor & Pylyshyn 1988; Braine & O'Brien 1998; Quilty-Dunn & Mandelbaum unpublished a; unpublished b). For example, Reverberi et al. (2012) presented

subjects with semantically odd but syntactically well-formed conditionals like "If there is a 3 then there is an 8" and then subliminally presented subjects with either the antecedent or the consequent (e.g., a 3 or an 8). They found that subliminal presentation of the antecedent facilitated the consequent but subliminal presentation of the consequent did not facilitate the antecedent. Effects like this can be explained by positing a mental logic that includes modus ponens but not affirmation of the consequent and pertains to the syntactic structures of propositional thoughts (Braine & O'Brien 1998; Quilty-Dunn & Mandelbaum unpublished b).

Many other effects of concepts and beliefs can be explained by syntactic structure (Quilty-Dunn & Mandelbaum unpublished a). One (rarely discussed) example is semantic priming. In a lexical decision task, for example, subjects have to discriminate a series of words and non-words (i.e., indicate whether 'bread' is a word [yes] or 'drabe' is a word [no]). If you read the word 'doctor' then you will later be quicker at discriminating a word like 'nurse' than you will be at discriminating words that are not semantically related to doctor (Meyer & Schvaneveldt 1971; McNamara 2005). Why should this be the case if all representations are unstructured vehicles for propositionally structured Fregean senses? The effect does not seem to go by way of propositional content since neither 'doctor' nor 'nurse' expresses a proposition. The explanation that thinking a thought like DOCTORS ARE OFTEN FOUND NEAR NURSES is necessary for every single instance of semantic priming is *ad hoc* and implausible. And even then, why should different unstructured vehicles that share only their abstract Fregean contents all facilitate recognition of the word 'nurse'? Without some computational explanation in terms of vehicles it's hard to see how an extreme Fregean could offer a causal explanation of the effect.

A better explanation appeals to word-sized mental representations—i.e., concepts—and explains the effect by appeal to associative links between those concepts that need not go by way of propositions. Moreover, the fact that your propositional thoughts are linked to subpropositional stimuli (e.g., if you hear the word 'nurse' then you'll be more likely to think the thought DOCTORS MAKE A LOT OF MONEY than you otherwise would be) is explained by positing that these concepts compose into propositional structures such that activating NURSE associatively causes the activation of DOCTOR, which then composes into thoughts like DOCTORS MAKE A LOT OF MONEY.

These are just a few examples of the explanatory benefits of taking mental representations to have syntactic as well as semantic structures. Extreme Fregeanism does not have this sort of explanatory power and should therefore be rejected. It might be the case that different sorts of vehicles and their different functional roles bring along with them different Fregean contents—I intend to be neutral about whether the semantic values of complex mental representations are Fregean senses, Russellian propositions, sets of possible worlds, or something else. But the view that all the mental structure we need can be offloaded to semantics is not sufficient to explain the relevant data; we need syntactic structure.

2.3—Individuating vehicles. The foregoing is meant to show that we need a distinction between contents and vehicles and that vehicles have structural features and are individuated independently of their contents. Earlier I argued that vehicles can't be individuated entirely by how the creature is disposed to use them in mental processes. How should we individuate them?

At least two options are available: one could individuate them in terms of functionalist counterfactuals or in terms of psychologically primitive properties. My goal here is not to make

substantive contributions to the literature on this topic nor to solve the problems of providing non-semantic individuation criteria for mental representations. I do want to make use of the idea of non-semantically individuated vehicles in order to understand the outputs of perception (and explain the data mentioned in the discussion of extreme Fregeanism). However vehicles are individuated, the data arguably requires that there be some means of individuation. Rather than settle the question here I will merely point toward possible options that have been explored—and critiqued—by other philosophers.

One option is that mental representations should be understood in terms of true counterfactuals. For some way W of using a representation in a mental process that is constitutive of that representation, a creature might lack the disposition to use the representation in W because of some intervening factor, and yet it could still be true that *if the intervening factor were not present*, the creature would be able to use the representation in W. That is, it might be that the representation is buried so deeply that the creature lacks the dispositions that stereotypically go along with it, but the representation is still apt to figure in the relevant processes despite the fact that it cannot.[6] On this kind of functionalist theory, the individuation conditions of a

---

[6] One might argue that dispositions are mutable in this way (Lewis 1997). But in the dispute between dispositionalists and representationalists in cognitive science, dispositionalism is not typically understood to allow dispositions to figure in processes to be buried this deeply. Eric Schwitzgebel, for example, argues that his dispositionalist view of belief rules out cases where beliefs are buried so deep in the mind that they do not occurrently dispose the believer to use the belief in certain ways (2013, 83). What matters is not that representations can fail to actualize their dispositional properties but rather that their stereotypical dispositions are completely buried due to extrinsic architectural facts. The version of dispositionalism critiqued above is thus, like Schwitzgebel's, a "superficial" dispositionalism (which I marked by calling representations on this view "mere dispositions"), while the attenuated and "deep" version of dispositionalism that allows for these sorts of counterfactuals to suffice for a mental representation to be tokened are compatible with the arguments made in this chapter. What matters is that token representations are not *mere* dispositions; perhaps certain deep dispositional properties can suffice for token representations *qua* mental particulars, a possibility I leave open (cf. Peacocke 2004).

representation are metaphysically tied to its being usable in W, but this disposition need not be actually present in order for the token representation to maintain its identity. This functionalist theory thus seems to avoid claiming that representations are mere dispositions while still allowing that the processes a representation is apt to figure in play a role in individuating its type.

A problem for this view is that functional roles might seem to change over time. Aydede (1998) and Prinz (2002) point this out, arguing that "functional roles are ineluctably holistic" (Prinz 2002, 97), in which case two mental representations in two people's minds or in the same mind over time cannot be type identical. But this objection could be avoided if there is a principled way of hiving off the right functional role and considering it from a counterfactual rather than actual perspective. For example, it might be that what individuates a type of mental representation is the kind of role that it would have independently of performance errors, abstracting away from changes in architecture and other representations housed in the same mind.

Much of the functional differences between my concept DOG and yours are arguably due to differences in our beliefs about dogs and the associations we have with dogs (e.g., with images and episodic memories as well as with other concepts). It does not seem *ad hoc* to control for differences in functional role that arise from relations to other stored representations. The individuation conditions for types of vehicles are an aspect of the architecture. Pylyshyn's (1984) influential characterization of architecture as the aspects of cognition that remain constant despite changes in the information stored in the mind would thus seem to legitimize the move of abstracting away from other particular representations in characterizing the functional individuation conditions for vehicles.

Suppose moreover that, were I to come to have all the same beliefs about dogs and associations with dogs, then the vehicle for my concept DOG would have the same functional role in my mind that yours does in your mind. The truth of this functionalist counterfactual may be sufficient for the mental representations to be tokens of the same type of vehicle even when we don't actually have all the same beliefs, associations, and dispositions. One would have to also abstract from other architectural variations, e.g., differences in working memory and long-term memory capacity, from performance errors, and perhaps other *ceteris paribus*-violating properties as well. But it is not obvious that some such functionalist counterfactual story would inevitably fail to individuate vehicles without succumbing to Aydede and Prinz's holism objection. Or perhaps there is some privileged set of computational processes that functionally individuate symbols (Schneider 2009).

Another option entirely is that mental representations should be understood as psychologically primitive. That is, there is some property that is not explicable in psychofunctional terms that marks each type of mental representation. In response to the question of what distinguishes two atomic representations, Fodor (2008) writes:

> I don't care. Type distinctions between tokens of primitive mental representations can be distinguished by anything at all, so long as the difference between them is of a kind to which mental processes are responsive. Since, by definition, basic representations don't have structures, type identities and differences among primitive Mentalese tokens are bedrock from the computational point of view. Tokens of primitive Mentalese formulas are of different types when they differ in the (presumably physical) properties to which mental processes are sensitive. Perhaps it's how long it lasts that determines whether a brain event is a tokening of [one concept] rather than of [another concept]. (I doubt that it is; but given the utter lack of a neuroscience of Mentalese, that guess seems about as good as any.)
>
> (Fodor 2008, 79)

Fodor casts the view in neural terms, which is problematic given the lack of any plausible neural candidates for individuating concepts (a problem to which "I don't care" is a less than satisfying answer). But the primitivist view need not be cast in neural terms. It is arguably plausible that there are levels of computation that mediate between the level of concepts and the level of bare neural activity (Dennett 1978; Lycan 1987; Craver 2007). In this case, the individuation conditions of representations that are primitive at one level might be explicable in terms of representations and processes occurring at the next level down. Types of representations would therefore be individuated in neural terms only at the lowest level of psycho-computational reality.

There are several advantages to this version of the primitivist view about individuating representations. One is that it comports with the existence of intervening levels. Another advantage is that it allows for representations to be multiply realizable. Casting individuation conditions in exclusively neural terms rules out the possibility of token representations in non-neural systems being type identical to representations realized in human brains.[7] Positing a multiplicity of psycho-computational levels allows for representations at the level of propositional attitudes (e.g., concepts) to be type-individuated by computational facts at the next level down; since this next level is not characterized in neural terms, it could be implemented in another physical substrate and thus allow two representations realized in neural and non-neural systems, respectively, to be tokens of the same type.

---

[7] Fodor says later that mental representations can be multiply realizable (2008, 90ff), but as Prinz (2011) points out, this move just asserts that there is a solution to the problem without actually solving it. And in any case, Fodor does not explicitly appeal to a multiplicity of computational levels as I'm doing here.

Here is a toy example. It might be that neural activity implements a connectionist network and that the connectionist network implements a language of thought architecture (Fodor & Pylyshyn 1988; cf. Smolensky 1990). This scenario so described contains three levels of representational ontology: a neural level, a connectionist level, and a symbolic language-of-thought level. Distinct nodes in the connectionist network might be realized by types of neural activations and some symbol in the language of thought might be realized by a certain pattern of weighted connections in the network. A connectionist network with the same weighted connections might be realized in a non-neural system with distinct physical properties individuating nodes, but with the same patterns of weighted connections between those nodes. Though the primitive representations in these two connectionist networks would be type distinct (since one would and the other would not be characterized in neural terms) they both implement the same symbol in the language of thought due to the functional equivalence of their weighted connections.

As long as there are relations of functional equivalence across systems with distinct physical realizations (and thus across distinct physically individuated primitive representations), those systems can in principle implement type-identical representations at a higher level of computational reality. It is crucial to the coherence of this story that the functional equivalence of these systems (and *a fortiori* the identity conditions for primitive vehicles at a higher level of computational reality) can be stated in terms that abstract away from the identity conditions of their primitive representations. Fortunately, this does not seem impossible. A connectionist network implemented in a brain might consist of n nodes with weighted connections between

them. A connectionist network with a completely different physical implementation might also consist of *n* nodes with the same weightings. What individuates nodes in the first network are neural properties; thus what makes it the case that the same node persists despite changes in weights will be the persistence of those neural properties. The individuation and persistence conditions of primitive representations in the second network will be different and so to that extent the two networks are type distinct—in other words, the difference is not *merely* a difference in implementation but is also a difference in the primitive symbols that figure in the two systems.

The systems nonetheless share properties characterized in the vocabulary of connectionist network, viz., the number of nodes and the weights between them, which are thereby multiply realizable properties. As long as it is the latter sort of property that is responsible for the individuation and persistence conditions of symbols at a higher level, then the two networks can implement type-identical symbols at a higher level (e.g., in the language of thought) despite differing in both their physical implementation and in the individuation and persistence conditions of symbols or nodes at the connectionist level. A full characterization of the two networks will not regard them as wholly type identical given the difference in primitive representations but the fact that they share some functional properties allows them to implement type-identical primitive representations at a higher level.

In a review of Fodor (2008), Prinz (2011) points out the implausibility of finding neural individuation conditions for vehicles. He also points out the lack of a clear and substantive proposal for non-neural individuation conditions of vehicles. The view just sketched above surely

does not constitute a detailed theory. Nonetheless, it points the way forward for a detailed theory, the development of which lies far beyond the scope of this dissertation.[8]

This section has raised more questions than it has answered. Moreover, I have not selected either of these solutions over the other (or ruled out other solutions), in part because they are too programmatic and the empirical details will likely play a decisive role in how we individuate vehicles. I confess to doubts about the coherence of both proposals, the former because of its potential for representations to collapse into mere dispositions and the latter because of its failure to provide a substantive characterization of the individuation conditions for mental representations. The latter seems less open to devastating, in-principle objections. It also allows for the total priority of representations to computational processes at the same level, which is preferable for providing substantive causal-mechanical explanation. I thus tentatively prefer the view that mental representations are psychologically primitive and should be cashed out in terms of multiply realizable functional properties of computations at a lower level (except at the lowest level, at which point vehicles are cashed out in neural or other non-representational terms).

For the purposes of this dissertation, however, what matters is that mental representations have rich syntactic as well as semantic properties. One of the outstanding goals of philosophy of

---

[8] It is worth pointing out that Prinz's (2011) claim that concepts comprise a wide array of information used for categorization including prototypes, exemplars, and theories seems to face a similar problem: what are the individuation conditions for the vehicles that make up the prototypes and other representations that constitute a concept? Any theory that appeals to mental representations arguably requires a vehicle–content distinction, and thus requires non-semantic individuation conditions for vehicles. One could invoke neurofunctionalism and empiricism and say that types of neural activations in perceptual cortical areas individuate types of vehicles—a constellation of views that comes with its own costs, of which the lack of multiple realizability is only one. The problem Prinz isolates is particularly pressing for Fodor given Fodor's claim that the syntax of concepts is proprietary and amodal. But anybody who thinks the vehicles of mental representation are both locally individuated and multiply realizable faces some version of this problem.

perception is therefore to provide a syntax and semantics for perceptual representation. The bulk of this dissertation is a contribution to this project. Little in what follows will hang on any particular view about the individuation conditions of mental representations. Before moving on to the syntax and semantics of perceptual representation in the following chapters, however, it will be helpful to clarify more foundational questions about mental architecture.

§3. The Representational Strategy, or: How Perception Talks to Cognition

3.1—The representational strategy and its translation problem. Much philosophical inquiry on the relation between perception and cognition concerns whether cognition affects perception, and if so, how. A comparatively neglected and arguably more central question is how perception affects cognition. It is as obvious a truism as any other that what we see affects what we think. A completed cognitive science should therefore be able to explain how perception can deliver information to cognition that allows us to update our beliefs and rationally plan action quickly and systematically.

An account of how perception talks to cognition needs to be informed by a theory of the syntactic and semantic properties of perceptual representations. If general anti-pluralism is true and all mental representations are couched in the same code, then perception delivers the very same sorts of representations that are used in thought (Pylyshyn 1984; Prinz 2002). In that case, whatever processes cognition uses to operate on perceptual representations may be general processes used in cognition to operate on cognitive representations. The question of how perception talks to cognition is thus the same as the question of how cognition talks to itself, since

perception and cognition speak the same language. There is, on this view, no special problem of how perception feeds into cognition so easily.

If instead one pursues the representational strategy for distinguishing perception from cognition then this problem cannot be so easily discharged. The representational strategy posits a difference in the kinds of representations used in perception and cognition. The view thus entails that perceptual processes output representations that are in some way nonconceptual and therefore distinct from representations used in cognition. This gives rise to a *translation problem*. If perception does not speak the same language as cognition, it can't talk directly to cognition; there must be some intermediating process proprietary to cognition that takes nonconceptual percepts as inputs and delivers conceptualized states as outputs (Mandelbaum forthcoming). Only once the translation process has finished can cognition act upon the (now syntactically reformulated) information delivered by perception.

All theories require a notion of sensory transduction, whereby physical magnitudes are transformed into syntactic magnitudes that can function as inputs to perceptual processes (see Pylyshyn 1984, Chapter 6 for an illuminating discussion). Though proponents of the representational strategy rarely acknowledge it, their approach also posits a sort of *cognitive* transduction whereby percepts are transformed into representations usable by cognition. Fodor foresaw this problem, writing that the function of perceptual modules is to "provide the central machine with information about the world; information expressed by mental symbols in whatever

format cognitive processes demand of the representations they apply to" (1983, 39). On Fodor's architectural approach, there is no need for a second layer of transduction.[9]

The severity of the translation problem depends on just how different percepts are from concepts. A greater degree of difference requires a greater amount of work to be done by the translation mechanism, and thus a more severe translation problem. We therefore need to know precisely what it means to say the outputs of perception are nonconceptual.[10]

3.2—Nonconceptual content and nonconceptual vehicles. The past two decades have generated a vast and stultifying literature on what is known as "nonconceptual content." The first clear formulation of the thesis that there is nonconceptual content in perception is often credited to Gareth Evans (1982). Nonconceptual content is supposed to be a type of content that differs from the content of cognition. However, what this means precisely has often been unclear, typically because of a lack of clarity concerning the vehicle–content distinction.

Richard Heck (2000) distinguished between the "state" and "content" versions of the appeal to nonconceptual content. The state version holds that the vehicles of perceptual representation are distinct from the vehicles of cognitive representation. The content version holds that the kind

---

[9] This discussion runs over a fine point of Fodor exegesis, namely that Fodor (1983) does not identify perceptual systems with modules. But this addendum relies on two almost entirely unargued-for assumptions: that perception eventuates in the fixation of belief and that the fixation of belief is unencapsulated from background knowledge. I think the former is not a universal truth since some perceptual processes deliver unconceptualized icons and the latter is in fact false—see Quilty-Dunn 2015. Fodor should therefore have counted modules as perceptual systems even when they fail to deliver beliefs and he should have also counted some processes that output perceptual beliefs as modular.

[10] I should add that the pluralist and anti-globalist architectural strategy I commend does not entirely avoid the translation problem. As will become clear in the next chapter, I think some perceptual processes output icons which are then conceptualized post-perceptually; the translation problem therefore arises for my view as well. However, there is reason to think that only early perceptual processes output icons while a great deal of the content of perception is couched in a discursive conceptualized format. The translation problem for a perceptual pluralist view is therefore far less dramatic than the one faced by proponents of the representational strategy.

of content that occurs in perception cannot be the content of a cognitive state like a belief, where the content of belief is often fleshed out in terms of structured propositions. I will use "nonconceptual content" to refer to content rather than vehicles (a practice which is unfortunately not uniform in the literature).

The nonconceptual content thesis has been defended by authors such as Heck (2000), Peacocke (2001), and Burge (2010a). It has been challenged by authors such as Speaks (2005) and Byrne (2005), who wonders how any content could not possibly be the content of a belief. I will not jump into this debate (though see Chapter 4). I will instead note that the question of whether perception has a distinct type of content is not obviously relevant to mental architecture in general or the translation problem in particular.

For example, many who accept that the *content* of perception is conceptual may nonetheless accept that the *vehicles* of perception are not concepts. They may hold that the content of both perception and cognition is a Russellian proposition (e.g., perhaps, Byrne 2005) or a set of possible worlds (Stalnaker 1998—though, in a testament to how confused and tacitly pro-structured-propositions this debate is, Stalnaker points out that given extant formulations of conceptual content, his view of propositions entails that both perception and cognition have *nonconceptual* content). However, commonality at the level of abstract contents does not solve the translation problem if the vehicles of that content have different general types of syntactic structures, i.e., different formats (see Chapter 2). If central cognition computationally transforms representations in response to their conceptualized (and perhaps propositional) syntactic structures, then a nonconceptual vehicle has to be translated into the proper conceptualized format

whether or not the input and output of the translation process share the same content (e.g., the same Russellian proposition).

Conversely, it does not seem impossible *a priori* that a vehicle in the format used by central cognition could carry nonconceptual content. Perhaps only concepts that are deployed in a stimulus independent fashion could count as having conceptual content (Camp 2009—though see Chapter 3, SEEING OBJECTS: A CASE STUDY IN THE UNDER-INTELLECTUALIZATION OF PERCEPTION). In that case, the visual system might in principle output representations that are syntactically type identical to representations used in cognition despite having different sorts of content. This view would avoid the translation problem, since computational processes in central cognition could operate on the outputs of perception just as well as on cognitive representations despite the difference in content.

From the perspective of mental architecture, therefore, it seems that the *vehicles* of perception are what matter, not what sorts of abstract *contents* are instantiated in those vehicles. That is not to say that the nonconceptual debate is not important for other reasons, e.g., for epistemology, phenomenology, or philosophy of language. Moreover, it is possible that a difference in format and functional role brings along a difference in content (Block unpublished). In any case, I will primarily consider the representational strategy insofar as it commits to type-distinct vehicles in perception and cognition.

Chapter 2 will contain a critique of the representational strategy so considered. While there are conceivably a variety of ways one might hold that the vehicles of perceptual representation are distinct from the vehicles of thought, the most prominent contemporary approach holds that the

format of perceptual representation is iconic or image-like while the format of thought is discursive or language-like (Carey 2009; Burge 2010a; 2014a; 2014b; Block 2014; unpublished). The next chapter will provide a critical evaluation of this view.

Before considering this view, however, I will sketch what I see as the most promising version of the architectural strategy.

§4. The Architectural Strategy

Perception cannot, I will argue, be distinguished from cognition by appeal to representational format.

Some architectural theorists distinguish perception from cognition by endorsing a full-fledged modularity thesis according to which perceptual systems are wholly immune to any substantive cognitive influence (Firestone & Scholl 2016). The form of perceptual pluralism defended in this dissertation requires that the border between perception and cognition is diachronically porous such that learning processes can push concepts into perception. I will not provide a theory of how this learning takes place nor what timescale separates synchronic from diachronic effects. Nonetheless, Chapters 2 and 3 will provide strong evidence for diachronic learning that pushes concepts into perception. Thus a full-fledged modularity thesis that denies such diachronic effects will not be feasible.

Instead, I think the most promising version of the architectural strategy argues that perceptual processes are (i) dependent on stimulation and (ii) synchronically informationally encapsulated from cognition to some interesting degree.[11]

There are likely to be more fine-grained differences as well, e.g., differences in the algorithms that are implemented. For example, it is unlikely that the algorithms implemented in deriving shape from shading in the visual system are implemented anywhere else in the mind.

I am not providing any argument here that perception is in fact always stimulus bound and informationally encapsulated. Instead, I want to sketch one way the joint in nature between perception and cognition (see Block unpublished) may be understood in light of the truth of perceptual pluralism (and, therefore, the fecklessness of the representational strategy). I will however seek to clarify these criteria, especially informational encapsulation, in a way that makes this version of the architectural strategy more plausible and empirically tractable.

4.1—Stimulus dependence. Perception starts with the transduction of proximal stimulation and outputs representations of distal stimuli. Part of what intuitively marks perception is its responsiveness to sensory transduction, while paradigm cases of propositional thought (e.g., planning to move to a new apartment) are unbound from present stimulation. Stimulus dependence as I interpret it requires that perceptual processes cease when proximal stimulation

---

[11] Interestingly, these map onto the two of Fodor's original nine criteria for modularity that have been most discussed: domain specificity and informational encapsulation. Prinz (2006) points out that domain specificity is hard to articulate in a plausible way given that few perceptual processes operate directly on sensory transductions. Perhaps, however, a modularist could appeal to hierarchically nested modules, the outputs of modules at one level serving as inputs to the next. In any case, the notion of stimulus dependence is slightly different; what matters is that perceptual processing in non-pathological cases will cease when stimulation ceases, not that all perceptual processes operate on a proprietary input domain.

ceases, barring pathology (and controlling for the time that mediates the occurrence of proximal stimulation and the running of the relevant process).

Some argue that stimulus dependence is the essence of perception. In particular, some theorists have argued that concepts are deployable independently of stimulation and perceptual representations are not (Prinz 2002; Camp 2009; Burge 2010b). Since I will ultimately argue that concepts are deployed in perceptual systems, I will not take this line—and I will critique this view of concepts in Chapter 3. Moreover, I do not think that being dependent on stimulation is sufficient for perception, since there may be processes that are stimulus bound and post-perceptual (e.g., certain forms of categorization and demonstrative thought).

There are some cases of perception-like processing that are free from stimulation, notably hallucination and mental imagery. However, hallucination is pathological and thus not at issue. Mental imagery is not pathological, but it is also not perception—though it does share resources and some aspects of representational format, as I will argue in Chapter 2.

Bence Nanay (forthcoming) argues that much of online perception involves mental imagery. Nanay's notion of imagery is quite broad however, encompassing any processing that is not completely *determined* by stimulation such as filling in the blind spot and representing three-dimensional surfaces. This processing is still *dependent* on present stimulation in the sense used here, since it is controlled by stimulation and ceases if stimulation ceases.

There is also the question of how different perceptual modalities respond to different stimulation. Differences in types of stimulation alone (e.g., retinal stimulation) do not seem sufficient to distinguish the modalities given the existence of cross-modal effects (Block

unpublished). However, a theory of what distinguishes perception from cognition does not necessarily furnish us with a theory of what distinguishes each perceptual modality. The appeal to stimulus dependence here is not intended to make any claims about what distinguishes visual processing from processing in other perceptual modalities.

The notion of stimulus dependence is not particularly deep, but it is arguably required to distinguish perceptual systems from encapsulated systems in central cognition, such as the cheater-detection module (Cosmides et al. 2010). Distinguishing encapsulated perceptual systems from encapsulated cognitive systems is the sole theoretical reason for which I appeal to stimulus dependence. Informational encapsulation is the more substantive criterion, so I turn to it now.

4.2—Informational encapsulation. There is a prodigious amount of philosophical and empirical research into the topic of *cognitive penetration* (e.g., Pylyshyn 1984; 1999; Levin & Banaji 2006; MacPherson 2012; Stokes 2014; Briscoe 2015; Brogaard & Gatzia 2016; Firestone & Scholl 2016). The basic idea of cognitive penetration is enshrined in Pylyshyn's famous formulation:

> [I]f a system is cognitively penetrable then the function it computes is sensitive, in a semantically coherent way, to the organism's goals and beliefs, that is, it can be altered in a way that bears some logical relation to what the person knows[.]
> (Pylyshyn 1999, 343)

Fodor developed another influential notion, the notion of *informational encapsulation.* According to Fodor, processing in "input systems" (i.e., encapsulated perceptual systems) "does not have access to all of the information that the organism internally represents; there are restrictions upon the allocation of internally represented information to input processes" (1983, 69). I will discuss the relation between informational encapsulation and cognitive penetration

later—I don't think it's obvious that they are deeply related. For now, however, I will try to characterize encapsulation more fully.

Encapsulation is a matter of information access. Saying that System A is encapsulated from System B means that A is constrained (in some way to be spelled out later) from accessing information in B (Scholl 1997). The notion of information access is foundational to the computational theory of mind. Every computational system must have some information available to it. Computational systems transform inputs into outputs, and what determines the relevant input-output mappings is the information available to the system together with the algorithms it implements. What distinguishes computational processes from mere transductions is that the former do, and the latter don't, bring stored information to bear on their computational transformations such that the output is not merely a slavish registration of the input (e.g., Pylyshyn 1984, Chapter 6).[12]

As an illustrative example, take visual "shape from shading," a process that uses information about the shading of an object's surface to derive its three-dimensional shape (Ramachandran 1988). The system that implements this process seems to "assume" that light comes from above and slightly to the left (Sun & Perona 1998). Somehow, the information that light comes from above and to the left guides the mapping from shaded two-dimensional surfaces to three-dimensional shapes performed by the shape-from-shading system. Talk of the system

---

[12] Burge (2010a) appeals to a similar distinction between information registration and genuine representation, but also allows for computation to consist in (e.g.,) mere weighting or averaging of registrations and thus going beyond transduction without involving genuine representation (2010a, 424).

making an "assumption" is a metaphorical way of stating that a certain piece of information structures the relevant computations.

This structuring can take place either explicitly or implicitly. In the explicit case, a representation with the content <light comes from above and slightly to the left>—or, more likely, some much more precise specification—is stored in a manner that makes it accessible to the shape-from-shading system. When the system is fed its inputs, it accesses this representation (and other stored information) and derives its outputs in accordance with whatever algorithms it implements. The notion of access involves the activation of a token representation and its functioning directly and without independent intervention in the relevant processes; a representation that is accessed by some process is computed over by that process.

In the implicit case, on the other hand, there's no such representation stored anywhere. Instead, the algorithms implemented by the system accord with the principle that light comes from above and to the left. That information isn't available for use by any process; it's merely implicit (or "embodied"—e.g., Devitt 2006, 45) in the operations of the system (Shea 2014). For this piece of information to be "accessed" is simply for it to accurately describe the transformations demanded by the relevant algorithms.

Note that 'explicit' in this context does not mean 'conscious'. Some (e.g., Fodor 1981b) argue that rules of grammar are explicitly stored in the language faculty and are accessed in deriving the syntactic properties of linguistic expressions, but would nonetheless deny that these representations are ever conscious. The present use of 'explicit' thus differs from the use of that term in discussions of implicit vs. explicit attitudes, such as the distinction between implicit biases

and conscious attitudes. On a view like Mandelbaum's (2016), for instance, implicit biases are explicit in the sense that they are realized by mental representations that are directly available for use in various mental processes, and are implicit only in the (very different) sense that they are unconscious.[13]

We can loosely use the term 'proprietary store of information' to describe both the explicitly and implicitly stored information used by a system to structure its input-output mappings. This term is loose in that implicitly stored information, unlike explicitly stored information, is not literally stored anywhere. Similarly, implicitly stored information, unlike explicitly stored information, is not literally accessed while running computations within the relevant system. Nonetheless, it's part of the set of information that shapes the processing of a given system, and can thus be counted as part of the proprietary store of information.

It's useful to have a neutral way of talking given how routinely difficult it is to determine whether a piece of information is stored explicitly or implicitly. The case of light coming from above and to the left demonstrates this methodological problem. Is that information an explicit or implicit part of shape-from-shading processing? This is not a rhetorical question, but it's difficult to know how to go about answering it. Another illustrative case is generative grammar (e.g., Chomsky 1980; Devitt 2006; Pietroski 2008). Critics notwithstanding, generative linguists believe

---

[13] Though 'implicit' in the sense of being unconscious does not entail 'implicit' in the sense of being unavailable for use in any mental process, the latter will entail the former given the reasonable assumption that only token representations can be conscious. Any token representation fails to be implicit in the latter sense, and thus the scope of explicit (in the sense of conscious) representation is limited to explicit (in the sense of available) representation.

that the rules of grammar are part of the language faculty's proprietary store of information. There is, however, no consensus on whether this information is explicitly stored.

With all this in the background, we can give a more precise characterization of encapsulation: System A is encapsulated from System B when A's proprietary store of information excludes information stored in B. While operating on its inputs to produce its outputs, A can only access a certain limited domain of information, and a piece of information held in B—even if it's semantically highly relevant to the processing in A—simply fails to fall within that domain, and thus cannot be accessed by A. Encapsulation in this sense underwrites the modularist explanation of recalcitrant perceptual illusions: despite the fact that a piece of information in cognition is semantically relevant to the illusory percept (e.g., one knows that the stimulus is illusory), that information falls outside the proprietary store of information available to perception, and thus the illusion persists.

A primary reason encapsulation matters for mental architecture is that it provides a good reason for positing distinct systems. If two processes occur in the same system, then semantically relevant information that is accessible to one process should typically be accessible to the other as well. For example, guessing where a person is from and deciding whether to ask them for directions are distinct cognitive processes, but a single piece information (e.g., that they only speak Finnish) can be relevant to both processes. In this example, when told that the person speaks Finnish you might form only a single representation, that person only speaks finnish, which is stored in central cognition and is accessible to both processes (e.g., it might lead you to guess that they are from Finland and, if you don't speak Finnish, to decide not to ask them for directions). This sort of case

would provide a compelling reason to think that the processes of guessing where a person is from and deciding whether to ask them for directions both take place in the same central-cognitive system.

If, on the other hand, two processes cannot access the same store of information, then there is good reason to think they don't take place in the same system. It's for this reason that Firestone & Scholl, for example, take the idea of widespread violations of the encapsulation of perception from cognition to threaten the idea of "a salient 'joint' between perception and cognition" (2016). Some theorists on the other side of the debate agree. Clark, for example, writes that the putative prevalence of violations of encapsulation "makes the lines between perception and cognition fuzzy, perhaps even vanishing" (2013, 190). Others, such as Block (unpublished), argue that violation of encapsulation is widespread but that perception is nonetheless distinct from cognition, appealing to factors like differences in representational format.

There has been a sustained attack on the thesis that perception is encapsulated from cognition in the experimental and philosophical literature (Prinz 2006; unpublished; Proffitt 2009; Balcetis & Dunning 2010; MacPherson 2012; Clark 2013; Mole 2015; Lupyan 2015; 2016; Block unpublished), along with replies from proponents of encapsulation (Durgin et al. 2009; Deroy 2013; Firestone & Scholl 2016).

Much of the debate has concerned nitty-gritty details of experimental design (Durgin et al. 2009; Proffitt 2009; Gantman & van Bavel 2014; Firestone & Scholl 2014). I will not jump into this aspect of the debate, partly because of the sheer size of the literature and partly because Chaz

Firestone and Brian Scholl (2016) have usefully pointed out a series of methodological flaws that make much of the extant experimental literature hard to evaluate.

Instead, I will discuss a more *a priori* point of contention. Many theorists (Lupyan 2015; Clark 2016; Block unpublished; Prinz unpublished) have claimed that attention can be a vehicle for widespread violations of encapsulation. Even some who are otherwise committed to encapsulation have expressed sympathy for this objection, including Firestone and Scholl (2016).

I will instead argue for an extreme thesis: no effect of attention, either on the inputs to perceptual processing or on perceptual processing itself, constitutes a violation of the encapsulation of perception.

§5. Attention, Encapsulation, and Impenetrability

5.1—Attention and perception. Attention is a form of modulation of perceptual processing that is typically taken to involve the selection or prioritization of information (though how this should be understood is a matter of debate—see, e.g., Nanay 2010; Mole 2011; Prinz 2012; Wu 2014). Attention can consist in some information being processed rather than other available information, or in constrained forms of modulation, such as increasing "gain," i.e., signal strength, and "tuning," i.e., noise reduction (Ling et al. 2009). Attentional modulation can increase spatial resolution (Carrasco & Barbot 2014), boost perceived contrast (Carrasco et al. 2004), and warp perceived distance (Vickery & Chun 2010).

In the visual system, attention comes in three broad forms. *Spatial* attention selects information from certain spatial locations and inhibits other locations. Spatial attention can be

overt, as when you move your eyes to change your gaze, or covert, as when you maintain fixation on a single point but attend to something in your periphery. *Feature-based* attention selects representations of certain features and inhibits others; for example, you might search for your favorite red shirt in your drawer and attend to red over the other available colors. Finally, *object-based* attention selects representations of certain objects; for example, you might watch a flock of birds and attend to a particular bird, tracking it as it moves.

Cognition can direct all three varieties of visual attention. For instance, you might attend to a certain location because you expect something to appear there; you might attend to a certain feature because you want to find your shirt; and you might attend to a certain object because it's a bird you intend to track. One might be tempted to conclude that these are all instances of cognition violating the encapsulation of perception.

Firestone and Scholl (2016) argue instead that at least some forms of attention merely constitute changing of inputs to otherwise encapsulated perceptual systems. An example would be when you choose to change where your eyes are fixated; though this is a cognition-driven change in perception, it operates not by violating encapsulation but instead by moving the eyes to change the inputs to perception. Though covert spatial attention doesn't literally involve moving the eyes, it's intuitively the same sort of effect and thus changes perception only by changing inputs (see also Deroy 2013).

Macpherson (2012) raises the possibility that feature-based attention directed by cognition seems to violate encapsulation (see also Mole 2015; Block unpublished). Unlike spatial attention, it's not simply a change in where we are looking, and it seems moreover to require antecedent

processing (e.g., it seems to require processing of red in order to attend to the red things in the room). A modularist may reply that attention does not only select *sensory* inputs, but also selects *representations* as inputs for downstream processing, even if those representations were the result of earlier stages in a hierarchy of modular processing.

Others argue that attention has widespread effects on perceptual processing that are not limited to selection and modulation of inputs (e.g., Mole 2015; Lupyan 2016; Clark 2016). Even Firestone and Scholl only take their argument to save encapsulation from "peripheral" forms of attention, and say they are "sympathetic" to an attack on encapsulation that focuses on non-peripheral forms of attention. They object only that such proposals have at present remained too abstract relative to the experimental evidence, thus leaving encapsulation-based theories of perception vulnerable to detailed accounts of non-peripheral forms of top-down attention.

5.2.—Attention and encapsulation. A methodological problem arises here: how do we determine which attentional effects might, and which might not, violate encapsulation? The debate, so stated, threatens to become about warring intuitions. The idea that covert spatial attention fails to violate encapsulation seems to rest on an intuitive parallel with overt spatial attention (e.g., Macpherson 2012, 28). Theorists like Macpherson (2012), Prinz (unpublished), and Block (unpublished) invite us to have the opposite intuitions about feature-based and (in Prinz's case) object-based attention, since these forms of attention seem to involve a tighter connection between the contents of our intentions and the contents of our resultant perceptual states. Indeed, Pylyshyn's classic formulation of cognitive penetration is of a "semantically coherent" causal relation between cognition and perception (1999, 342), and the notion of semantic coherence is hard to explicate

without intuitions. For example, feature-based attention to red things that alters perception of redness in response to thoughts about redness seems to instantiate a causal, semantically coherent relation (MacPherson 2012; Block unpublished).

Pylyshyn's influential formulation of cognitive penetration in terms of semantically coherent causal relations between thought and perception thus threatens to make mundane forms of attention into examples of cognitive penetration. Note, however, that the presence of a semantically coherent causal relation between a thought and the output of some perceptual process says nothing about stores of information. For all that Pylyshyn's informal definition of cognitive penetration says, there could be semantically coherent causal relations between thoughts and percepts that don't involve a change in the proprietary store of information available to the relevant perceptual process.

Semantically coherent effects mediated by attention would constitute just this sort of relation. If a representation is held in a system's proprietary store, it's directly available to be accessed in running computations. No additional process is necessary for it to affect processing. Every computational system can immediately access its proprietary information store either implicitly—i.e., running algorithms that accord with it—or explicitly—i.e., activating the representation and having it enter into the online computational process. Access of proprietary information is direct or immediate in the sense that there's no intervening process except the activation of the stored representation.

Consider, for example, the access of a belief stored in long-term memory in central cognition. Barring some sort of malfunction or independent intervention, once the belief is

accessed (which, since it's explicitly stored, amounts to activating the representation), it will enter into central cognitive processes like inference (Braine & O'Brien 1998) and facilitating lexical recognition (Meyer & Schvaneveldt 1971). The fact that representations stored in a system currently running a computation are immediately entered into computations *via* access explains why, to reuse an example mentioned earlier, subliminally presenting the minor premise in a *modus ponens* argument after consciously presenting the major premise results in acquisition of the conclusion (Reverberi et al. 2012). Assuming that *modus ponens* describes an algorithm implemented in central cognition, then once the major premise (e.g., 'If there's an X then there's a Y') is represented in central cognition, activating the minor premise (e.g., 'There's an X') is causally sufficient to yield the conclusion as output (e.g., 'There's a Y'). The representation, once accessed, is integrated with currently running processes to yield an output.

It should be clear at this point why an effect of cognition on perception mediated by attention cannot violate encapsulation: it is not a form of information access. A state in cognition may cause an attentional effect on perceptual processing and thereby alter the output of perception. Moreover, it may do so in a semantically coherent way, such that the output of perception is semantically related to the state in cognition. But the state in cognition isn't thereby just another representation housed in the perceptual system's proprietary store of information. Instead, it can only exert an influence on the perceptual system by way of attentional selection and modulation. If a representation in cognition can only affect perceptual processing *via* attention, that immediately separates it from representations in the relevant proprietary information store; the latter are accessed directly, while the former isn't. States in cognition that can only affect perception

via attention are thus never part of the store of information proprietary to a perceptual system, and so don't violate the encapsulation of perception from cognition.[14]

Take Fig. 1, which to an innocent eye may appear to contain a chaos of dots. However, upon learning that the image depicts a Dalmatian, the dog may become visible. Suppose that this effect works through activating your concept of Dalmatians and using conceptual knowledge of what they look like to guide which locations and features you attend to. It's thus not the case that this cognitive information is *accessed* in the relevant perceptual processes. The cognitive states instead guide processes of spatial, feature-based, and, ultimately, object-based attention that modulate perceptual processing. The store of information proprietary to the relevant perceptual system is left intact.

[14] One can also think of information access as defined by the array of parameters that can be set in a computation. For example, distance perception may contain a parameter for whether the perceived object is desired (Balcetis & Dunning 2010)—e.g., if the object is desired, then the parameter is set to YES. Effects mediated by attention don't involve a parameter being set directly by the presence of a state in cognition, so such effects don't violate encapsulation.

Figure 1—Top-down attention at work

Similarly, Mole (2015) cites Kravitz and Behrmann (2011) as providing experimental evidence for top-down attention operating within (rather than before) perceptual processing. Kravitz and Behrmann found effects of attentional cues on a letter-discrimination task that were driven by the shape, color, and location of previewed objects. A cue would improve letter discrimination not only if the letter appeared at the same location or in the same object as the cue, but also if it appeared in the location of another object that shared features (e.g., shape or color) with the object in which the cue appeared. Mole (as well as Kravitz and Behrmann) argues that spatial, feature-based, and object-based attention interact to organize perceived scenes and guide perceptual discriminations. Mole argues further that this interaction can be sensitive to learned categories such as letter identities (2015, 230–231).

As mentioned above, a modularist could reply that these interactions of attentional effects take place at the output-input interfaces of hierarchically nested modules. But the point presently at issue is that encapsulation isn't violated *even if* attention operates within modular processing.

Attention is the selection and modulation of perceptual signals, not a form of information access. Selection and modulation within modules may, for example, affect which stored information is accessed in online perceptual processing by biasing competition between signals for "representation, analysis, or control" (Desimone & Duncan 1995, 194; see Mole 2015, 231ff). The cognitive states that drive such attentional effects, however, aren't thereby among the information accessed in the relevant perceptual processes.

One might object that attention simply *is* perceptual processing. Mole, for instance, writes that attention is "inherent to the basic structure of the neural architecture" of perception (2015, 233). Perhaps, in that case, cognitively driven attention is *ipso facto* a case where a perceptual process accesses cognition. But this claim, depending on how it's meant, is either trivial or false. If the claim is merely that attention (*qua* perceptual process) accesses information in cognition, then it's trivial; nobody on any side of the debate about encapsulation would deny that attention can be driven by (and hence can access) cognitive states. If, on the other hand, the claim is that the perceptual states and processes selected and modulated by attention thereby access information in cognition, then it's false. There is a clear asymmetry between (i) a process that accesses representation R and (ii) a process that's affected by a separate process that accesses R. My argument thus does not require arguing that attention is a separate system from perception, but only that attention is not identical to the processes (e.g., distance perception) it modulates. Since attention *modulates* such processes, it cannot be *identical to* such processes, even if it is genuinely part of the perceptual system and it deeply intertwined with the processes it modulates.

Consider a putative case where a cognitive state is genuinely accessed by perception. For example, suppose that a visual subsystem that implements perception of distance can access a desire for the perceived object and thereby make it appear closer (Balcetis & Dunning 2010). In that case, the store of information accessible to the distance-perception system includes desires stored in central cognition. Suppose instead, however, that the desire guides attention such that you attend more to things you want, and the perceptual signal is thus boosted (even, we can stipulate, within the relevant perceptual process rather than at the input stage). Boosting the signal then causes the object to appear larger and closer. In the second case, there's no explanatory reason to say the desire is part of the store of information accessible to the distance-perception system. The only "perceptual process" to which the desire is accessible is the deployment of attention.

It turns out that Firestone and Scholl's question of whether attention is merely a changing of inputs is largely beside the point. Even if top-down attention operates *within* the processing of a module, it doesn't violate the encapsulation of that module. The module can still retain a proprietary store of information that affects its outputs without being mediated by attention or any other mechanism (aside from information access). There can thus, in principle, be widespread effects of top-down attention on "every stage and level of processing" (Clark 2016) without violating the encapsulation of perception.

5.3—Encapsulation and impenetrability. Both proponents and detractors of cognitive penetration have tied the notion of cognitive impenetrability to the notion of informational encapsulation. However, as came out in the discussion of attention, the two notions are quite different. To say that an effect of cognition on perception constitutes cognitive penetration is to describe the

character of the effect, viz., that it is semantically coherent (Pylyshyn 1999). To say that an effect of cognition on perception constitutes a violation of informational encapsulation is to describe the mechanism of the effect, viz., that it is an instance of information access (Fodor 1983; Scholl 1997).

A state in cognition could in principle have a semantically coherent effect upon perceptual processing without being accessed by that processing. For example, a state in cognition might cause the formation of a mental image, and the mental image might alter the percept; every link in this causal chain respects the content of the representations involved, so the effect seems to be semantically coherent (MacPherson 2012). Nonetheless, it need not be the case that the perceptual process that is so affected actually *accesses* the cognitive state. Instead it seems that only the image is a candidate to be an actual constituent of the perceptual processing rather than the cognitive state. Attention, I have argued, can modulate perceptual processing without the perceptual process accessing cognitive states and thereby constituting a violation of encapsulation. It may therefore count as another kind of semantically coherent effect of cognitive on perception (i.e., cognitive penetration) that does not violate the encapsulation of perception.

Some proponents of cognitive impenetrability build in the requirement that cognitive penetration must be not only semantically coherent but *direct*, such that effects mediated by imagery and attention don't count (Gross 2017; Block unpublished). But why should directness matter? If what matters is that the character of the effect is semantically coherent, whether the effect is direct is irrelevant. If instead what matters is information access—that is, if informational encapsulation is at issue rather than cognitive impenetrability—then directness does matter. A piece of information is either computed by the perceptual process or it is not; if it is not, then it is

not accessed. This is the case whether or not that piece of information has an *effect* on the process, and whether or not that effect is semantically coherent.

I've argued that a cognitive state can penetrate perceptual processing without violating encapsulation. This dissociation also goes the other way—a cognitive state could in principle violate the encapsulation of a process without cognitively penetrating it. This possibility arises from the fact that informational encapsulation is a computational and therefore syntactically specifiable notion whereas cognitive penetration as defined by Pylyshyn is specified in terms of semantic relations between representations. Consider a process that can output any of an array of symbols $S_1...S_n$, including two symbols S_2 and S_3 that are semantically equivalent but are syntactically distinct—they might be in distinct formats, or the system might be built in some peculiar fashion such that it has a semantically redundant store of atomic symbols. Suppose that on a particular occasion, the input to this process causes it to output S_2. Suppose further that some state gets tokened in cognition and is thereupon accessed by the process, and computing over the cognitive state causes the system to output S_3 instead of S_2. This is a clear violation of encapsulation, since the state in cognition is accessed, computed over, and alters the output of the process. Nonetheless, since the content of the output was not altered, there was no semantically coherent effect and thus no cognitive penetration. This example may seem exotic, but it points out that cognitive penetration is in principle doubly dissociable from the violation of informational encapsulation.

For the purposes of mental architecture, informational encapsulation is arguably more to the point. If perceptual processes never access information stored in cognition, then they are totally

encapsulated from cognition even if cognition has massive impacts on perception that are not *via* information access. The question of whether cognition impacts perception in a semantically coherent way might be important for questions about epistemology and consciousness, but it is not obviously a foundational question for theories of mental architecture.

§6. Dimensions of Encapsulation

Before concluding this chapter and moving on to discuss the syntax and semantics of perceptual representation, it will help to clarify one more aspect of informational encapsulation. Whether perceptual systems are encapsulated seems like a yes-or-no question—either perception can access information in cognition or it cannot. Then again, writers in the literature on modularity often talk about modules as being encapsulated *to some degree* (e.g., Pacherie 2008; Burge 2010b; Shea 2014; Butterfill & Apperly 2016).

I have said I think that a promising version of the architectural strategy holds that perceptual systems are stimulus driven and synchronically encapsulated from cognition to some interesting degree. As far as I am aware, nobody has tried to provide a detailed taxonomy of the ways in which encapsulation can come in degrees as opposed to being a binary yes-or-no phenomenon. I will now try to sketch such a taxonomy. There is no method that underlies this taxonomy, but it does group somewhat into families of dimensions of encapsulation.

These dimensions do not necessarily commensurate with one another. That is, there may be no fact of the matter whether a process that is highly encapsulated along dimension A and weakly encapsulated along dimension B is more or less encapsulated than a process that is weakly

encapsulated along A and highly encapsulated along B. Moreover, I will refrain from drawing out the various ways in which these dimensions might combine (which they surely can). In many cases it is plausible that a system that is not perfectly encapsulated along one dimension will fail to be perfectly encapsulated along others.

A final caveat: as just argued, indirect effects and merely semantically coherent effects of cognition on perception do not constitute even a weak violation of encapsulation. Thus a taxonomy of graded dimensions of encapsulation should exclude such cases.

6.1—Number. One group of dimensions of encapsulation has to do with the amount of states or processes involved.

6.1.1—Number of processes. Perhaps the simplest grade of encapsulation is the number of perceptual processes involved. Philosophers sometimes talk of the entire visual system as a single module, but that is a misleading way of talking. Instead, the visual system comprises a plethora of distinct processes. Thus one way in which encapsulation might come in degrees is the number of perceptual processes that can access states in cognition. A perceptual system made up of 50 processes only three of which can access cognitive states is more encapsulated than a perceptual system made up of 50 processes nine of which can access cognitive states. This is *ceteris paribus*, of course, since the second system might be far more encapsulated along other dimensions—I will hereafter drop the *ceteris paribus* qualifications.

This dimension is strange, since encapsulation should typically be understood at the level of processes rather than systems; nonetheless, the tendency to ask whether *the* visual system is encapsulated makes this dimension useful nonetheless.

6.1.2—Number of parameters. With respect to a single process, the process could have more or fewer parameters set by accessing states in cognition. For example, one process might have a YES/NO parameter for whether the object is desired while another process might have that parameter as well as a parameter for whether the object is believed to be large. The second process is less encapsulated than the first because it has more parameters to be set by states in cognition.

6.1.3—Number of states. The mind might be constructed so that only a certain number of beliefs and desires are accessible to perception at one time. A perceptual process that can access a greater number of cognitive states is thereby less encapsulated than a process that can access fewer cognitive states. This dimension might be shaped by the types of cognitive states in question, which is mentioned below.

6.2—Strength. Another group of dimensions of encapsulation has to do with *strength*.

6.2.1—Strength threshold. It is plausible that beliefs vary in strength. For example, I believe that I live in Brooklyn and I believe that Lee Harvey Oswald killed JFK alone, but I believe the former more strongly than the latter. Strength is often equated with subjective probability (e.g., Pettigrew 2013), though there may be reason to doubt that strength *qua* recalcitrance correlates very well with strength *qua* subjective probability (Mandelbaum 2010).

Desires also seem to have degrees of strength as well. For example, one might want to complete one's dissertation and one might want to play video games, and the latter desire might be stronger than the former.

One dimension along which encapsulation might come in degrees is the strength required to affect perception. For example, a perceptual process might only access beliefs with strength .9

and up, while another process might access beliefs with strength .6 and up, and still another might access beliefs with strength .2 and up. The second process is less encapsulated than the first, and the third is even less encapsulated.

6.2.2—Strength sensitivity. Another way in which strength might grade encapsulation is the sensitivity of the perceptual process to the strength of the cognitive state. For example, take the Balcetis and Dunning (2010) experiment purporting to show that the process of perceiving distance computes over desires. The process of distance perception might have a simple YES/NO parameter for whether the object is desired; if the object is desired, then it modulates the perceived distance according to some set function (e.g., the object is represented as 5% closer). Or instead the process might access not only *whether* the object is desired, but also the *strength* of the desire, and modulate the perceived distance as a function of the strength of the desire. For example, if the object is weakly desired, it might be seen as 2% closer, and if it is strongly desired, it might be seen as 10% closer. This second process would be less encapsulated than the former since it not only accesses the cognitive state, it accesses and is modulated by the degree of strength of that state.

6.2.3—Strength of effect. A third way in which strength might grade encapsulation is the strength of the effect on the perceptual process. Here 'strength' is used somewhat differently. For example, the function by which distance perception is altered by the presence of a desire could be greater or lesser. If an object is desired, one process might represent the object as 1% closer while another process might represent it as 5% closer. The latter system is less encapsulated than the former since the degree to which the input-output function is altered in virtue of accessing the cognitive state is greater.

Note that this dimension is distinct from the number-of-parameters dimension, since a process with fewer parameters set by cognition might nevertheless have its computations altered to a greater degree than a process with more parameters set by cognition.

6.3—Type. Another group of dimensions of encapsulation concerns the types of states and processes involved.

6.3.1—Types of processes. At the level of the perceptual system, we can ask not only how many processes in that system are encapsulated, but also which *types* of processes are encapsulated. For example, early visual processes might be encapsulated while later ones are not (Pylyshyn 1999), or vice versa. Or perhaps perceptual processes that operate on one format are encapsulated while processes that operate on another format are not. This dimension is particularly vague, and it is not obvious how to render this a graded phenomenon, but it is nonetheless a dimension of variation of encapsulation at the level of systems. It may become more obviously graded in interacting with other dimensions, e.g., number of processes.

The rest of the dimensions in this category concern the types of cognitive states that are accessible to the process in question.

6.3.2—Types of attitudes. One process might only be able to access desires, while another could access hopes, intentions, and desires, and a third could access those states as well as beliefs. The third process is the least encapsulated, followed by the second and then first, because it has access to the most diverse array of cognitive states.

6.3.3—Types of formats. Cognitive states might vary in format. A perceptual process might be able to access states in every format contained in central cognition, while another might only be able to

access states in a few formats or only one format. This difference tracks a difference in degree of encapsulation.

6.3.4—Peripherality. Some states are more "peripheral" than others because they concern perceptible phenomena (Quine 1960). A process might be able to access peripheral beliefs like the belief that bananas are yellow but not more "central" beliefs like the belief that 3+4=7. A process that can access more central as well as more peripheral beliefs is thereby less encapsulated.

6.3.5—Conscious accessibility. Many putative examples of violations of encapsulation involve consciously inaccessible states (e.g., implicit biases), while many putative examples of encapsulation involve conscious beliefs (e.g., known illusions). A process that can access conscious as well as unconscious states is thereby less encapsulated than a process that can only access unconscious states.

6.4—Gateways. A fourth group of dimensions of encapsulation concerns what I will call "gateways." A mind might be constructed such that perception typically cannot access beliefs, say, unless the belief that p is tokened, in which case perception can thereupon access other beliefs. In this sort of case, the belief that p functions as a gateway, since it razes the border between systems and allows previously inaccessible states to become accessible. I am not sure whether any mind in fact works this way, but I include gateways here in the interest of mapping out the logical space.

6.4.1—Number of states needed. The gateway might be the belief that p, or it might be the belief that p and the belief that q. A process that requires more states to be present in order to facilitate access is more encapsulated since it is more restrictive about the conditions under which encapsulation can be violated.

6.4.2—Strength of states needed. Another gradation can be due to the strength needed. For example, a process that requires the belief that p and the belief that q both at strength .8 is more encapsulated than a process that only requires the beliefs to have strength .2 to facilitate access.

6.5—Summing up. This taxonomy is not intended to be anything like an exhaustive one. Nonetheless, it is important to understand how encapsulation might come in degrees, and these are 13 distinct dimensions along which a system or process can be more or less encapsulated. The hypothesis that perceptual processes are encapsulated from cognition "to some interesting degree" can be fleshed out along these (and perhaps other) dimensions.

§7. Conclusion

This chapter was intended to clarify some of the background issues concerning mental architecture and the vehicles of mental representation. The rest of the dissertation will concern the syntactic and semantic properties of the outputs of perception. I will defend the thesis that perception delivers multiple representational forms, including those used in thought. We thus cannot distinguish perception from cognition by appeal to differences in the vehicles deployed. In this chapter, I have tried to sketch an architectural strategy for distinguishing perception from cognition that appeals to stimulus dependence and some degree of synchronic encapsulation. This dissertation is not a defense of this strategy, which would require its own dissertation. Nonetheless, this chapter has provided some clarity on what exactly such a strategy would be committed to.

PERCEPTUAL PLURALISM

Chapter 2

"Discursive representations do, but iconic ones do not, impose principles of individuation on the domains in which they are interpreted. I don't want to talk about this at length because I'm scared to."
        —Jerry Fodor[15]

"Perceptual psychology assumes an iconic format for perceptual representation."
        —Tyler Burge[16]

Perceptual systems respond to proximal stimuli by forming mental representations of distal stimuli. A central goal for philosophy of perception, therefore, is to characterize the representations delivered by perceptual systems. It may be that all perceptual representations are in some way proprietarily perceptual, and differ from the representational format of thought. Or it may instead be that perception and cognition always trade in one common code. This chapter rejects both approaches in favor of *perceptual pluralism*, the thesis that perception delivers a multiplicity of representational formats, some proprietary and some shared with cognition.

---

[15] Fodor 2007, 109.

[16] Burge 2014b, 494.

§1. Introduction

Most philosophical writing on the nature of perceptual representation focuses on content. For example, philosophers have debated whether perception has conceptual content (McDowell 1994; Byrne 2005) or nonconceptual content (Evans 1982; Heck 2000; Peacocke 2001), high-level content (Bayne 2009; Siegel & Byrne forthcoming) or low-level content (Brogaard 2013), and rich conscious content (Dretske 2004; Block 2011) or sparse conscious content (Cohen & Dennett 2011), just to name a few disputes.

This preoccupation with content, though fruitful in some respects, is peculiar. Elsewhere in philosophy of cognitive science, it has been common to examine the *vehicles* of content—mental representations themselves, and how they, rather than their contents, are structured. For example, debates about whether there is a language of thought (Fodor 1975; Dennett 1978; Fodor & Pylyshyn 1988; Smolensky 1990), whether some thought is map-like (Camp 2007; Rescorla 2009a; Blumson 2015), representations in the language faculty (Chomsky 1980; Devitt 2006), and the nature of mental imagery (Pylyshyn 2003; Kosslyn et al. 2006) all primarily concern the structures of representational vehicles.

There is good reason for this focus on vehicles throughout cognitive science. Mental representations are the elements of mental computational processes. It is in virtue of their vehicular or "syntactic" properties rather than their contents or "semantic" properties (which may be such exotic entities as sets of worlds or Fregean senses) that representations have the particular computational roles they have in the systems in which they occur. An understanding of how

mental representations function in the mind and produce behavior thus requires an understanding of their (syntactic) structures.

In a discussion of cognitive maps as alternatives to propositional thoughts, Michael Rescorla laments that philosophical research on perception "generally emphasizes content rather than the vehicles of content. Participants in the debate seldom even mention mental representations" (2009a, 380). As a result of this focus on content, there is comparatively little philosophical discussion about what kinds of mental representations are constructed by the visual system.

Nonetheless, in recent years a movement has been building that seeks to characterize the structure of perceptual representation as a way of distinguishing perception from cognition. If the vehicles of perceptual representation (i.e., percepts) are structured differently than the vehicles of cognitive representation (i.e., thoughts), then we may be able to draw the perception–cognition border by appeal to such differences.

Susan Carey (2009, 8), Tyler Burge (2014a, 488; 2014b, 574), and Ned Block (2014, 560; unpublished) all argue that perceptual representations are to be individuated at least partly in virtue of their *format*. More specifically, they claim that the format of perceptual representation is iconic while the format of thought is discursive. This chapter is a critical evaluation of this thesis. I will argue instead for the pluralist thesis that perception delivers both iconic and discursive representations.

First I will clarify the notion of representational format and the distinction between iconic and discursive representations in Section 2. Then, in Section 3, I will survey evidence in favor of

iconic representation, arguing that there is defeasible but abductively valid reason to posit iconic representation in perception; this section will also function to empirically validate the notion of iconic representation outlined in Section 2. In Section 4, I will argue that object perception delivers representations in a discursive, non-iconic format. I will respond to various objections to the claim that perceptual object representations are discursive in Section 5. Finally, I will gesture toward other possible cases of non-iconic perceptual representation in Section 6 before concluding.

§2. Representational Format

2.1—How to think about format. The word 'bed' looks sort of like a bed. Thus the very same item can be used to pick out a bed in two different ways: linguistically and pictorially. For instance, suppose someone writes a short note whose contents are simply: 'I bought a new bed'. The typical interpretation would be that she's using 'bed' as a word, and so expressing the thought I BOUGHT A NEW BED in words. But suppose that in fact she has a peculiar lapse in memory and forgets the word for bed, deciding instead to write 'I bought a new' and then draw a picture of a bed (a picture that looks just like this: bed), so expressing her concept of a bed with a picture.

What is the difference between the linguistic item 'bed' and the picture 'bed'? The biggest semantic difference is that the pictorial use of 'bed' represents spatial properties and relations that aren't explicit in the linguistic use of 'bed'. But even a detailed linguistic description of all these features would not add up to a picture of a bed. So if the difference is not purely semantic, what is it?

The key difference consists in how the parts of the representational vehicle are used. In particular, the difference lies in the relations between parts of the representations and parts of the represented thing. Consider 'bed' as a picture. Parts of the picture represent parts of a bed: the vertical lines at the left and right ends represent a headboard and a footboard, and the tops of the curved lines represent the mattress (and the middle of the picture represents the middle of the mattress, the parts on either side of the middle of the picture represent the parts on either side of the middle of the mattress, etc.). Now consider 'bed' as a word. No part of the word represents part of a bed: the 'b' does not represent anything, let alone a headboard or part of a mattress.

This case provides a toy example of representational format in action. The difference between the pictorial and linguistic uses of 'bed' is not fundamentally a semantic one; rather the difference consists in the structure of the vehicle and how it is exploited.

Formats are general types of vehicular structures. The sentence 'This is a tiger' and a picture of a tiger differ in representational format. The phrases 'brown cow' and 'large yawning tiger' have distinct vehicular structures but do not differ in format, since they are both instances of English phrases—the difference in vehicular structure is not significant enough to mark a difference in format. Different formats are akin to different languages, such that representations that are couched in different formats typically cannot compose together (though new hybrid formats, like hybrid languages, can arise through convention). Examples of distinct representational formats include maps, graphs, diagrams, sentences, photographs, hieroglyphs, and blueprints. Some of these are more or less similar to one another, but they all differ enough in how they exploit the

structures of relevant representations and compose those representations that they each constitute a distinct format.

2.2—Iconic vs. discursive. Perhaps the most influential and fundamental distinction between formats in cognitive science is the distinction between iconic and discursive representations. According to the mainstream cognitivist viewpoint in the past half century, beliefs and other propositional attitudes are discursive (Fodor 1975; 1987; Chomsky 1980; Pylyshyn 1984; Prinz 2002; Carey 2009; Burge 2010). The systematic and productive ways that concepts seem to recombine into thoughts, the formal/logical character of deductive inference, and the word-sized associative links involved in semantic priming all call for a theory of propositional thought as literally structured like a sentence or proposition and breaking down into smaller parts that are syntactically concatenated with one another (Quilty-Dunn and Mandelbaum unpublished a; unpublished b). I will therefore assume that at least some thoughts are discursive.

Many who believe that thought is discursive also believe that there are mental icons, endorsing a pluralist view about mental representation generally (Fodor 1975; Kosslyn 1980; 1994; Dretske 1981; Kosslyn et al. 2006; Fodor 2008; Carey 2009; Rescorla 2009a; Burge 2010; Kulvicki 2015a; Block unpublished). Before surveying the empirical evidence in favor of iconic representations in perceptual systems, it will be useful to characterize the distinction between iconic and discursive formats.

There are two key differences between iconic and discursive representations.

The first is that parts of icons correspond to parts of what they represent, while this need not be true of discursive representations. Typically, there are also distance relations between parts of icons that correspond to distance relations in what is represented.

The second difference is that icons encode properties and individuals holistically, and not by means of a canonical decomposition into constituents that stand separately for distinct individuals and distinct features (see Green & Quilty-Dunn unpublished).

Consider the difference between a picture of Bob Dylan wearing a checkered shirt (Fig. 2) and the sentence

(1) Bob Dylan is wearing a checkered shirt.

Figure 2—Dylan wearing a checkered shirt

One can segment the picture arbitrarily and preserve semantic significance. While we may find it natural to draw a line around Dylan, the picture does not come pre-segmented in that fashion. Someone with different interests may segment the picture in line with Figure 3.
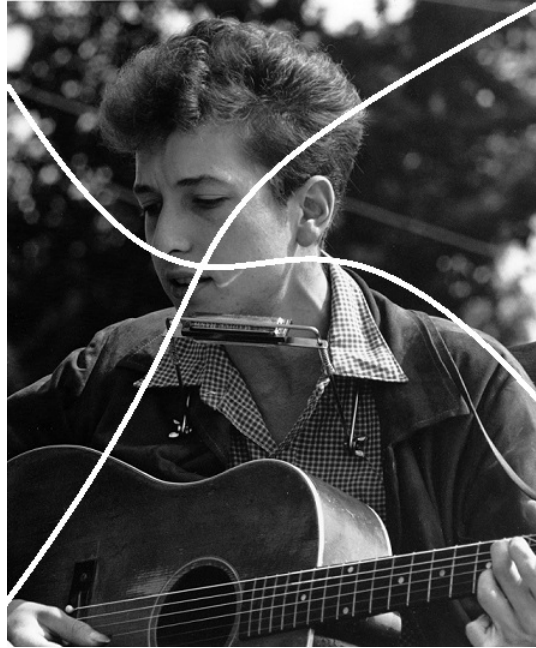
Figure 3—Dylan segmented

The segmented parts are each representations of some part of the scene, as is any other segmentable portion. This is why Stephen Kosslyn, for example, writes of mental images that "a 'part' can be defined arbitrarily, cutting up the representation in any way; no matter how you cut it, the part will still correspond to a part of the object" (1994, 5). Furthermore, distance relations between these parts track distance relations in the scene. Point simultaneously at the part that represents the top of Dylan's head and the part that represents his chin; now instead of pointing at his chin move downward to a part that represents his chest. There is now more distance between the two parts you're pointing at on the surface of the image, and this increase in distance tracks an increase in distance between the two represented parts of Dylan. Finally, any part of Figure 3 that you are pointing at represents not only some spatial part of Dylan (or of the tree behind him) but also values along color dimensions like lightness and dimensions of shape, size, and perhaps orientation as well.

As a contrast, the syntactic structure of (1) precludes free segmentation. (1) has a right way of carving it up, viz., into 'Bob Dylan', 'is', and 'wearing a checkered shirt', which themselves decompose down to individual words. One could segment out the part 'Dylan…wearing', but this is not a meaningful representation that composes according to some compositional semantics to generate the whole sentence. Or one could carve it up so that one of the parts is 'ob Dyla', which involves no meaningful representation at all. The sentence thus has at least one *canonical decomposition*, or right way of carving it up, and many wrong ways of carving it up. The parts that are individuated by a canonical decomposition are *constituents*. Thus discursive representations, unlike iconic representations, always either have constituents or are compositionally efficacious atomic representations like individual words or concepts. Icons have parts but typically lack constituents in this technical sense of the term (Fodor 2007).

Pure icons like Figure 2 (i.e., which do not have labels or any other explicit discursive apparatus amended to them) arguably always lack constituents; there are other hybrid icons, like maps and graphs, that do have constituents (Camp 2007; Rescorla 2009a, 2009b).

Distance relations among parts of (1) also do not track distance relations in what is represented. And finally, some parts of (1) represent individual items or properties without being holistically bound up with other items or properties. For example 'checkered' represents a surface texture without representing the object that has that texture, its shape, its location in the scene, etc.

One might object that many or all icons at least have *privileged* decompositions. For example, perhaps Rothko intended his shapes to be perceptually segmented separately from one another, and this means that it would be wrong to carve his paintings along other boundaries. It is

not obvious that in these cases we should read the privileged boundaries into the syntax of the icon rather than offloading it onto the discursive mental states of artists and consumers. Nonetheless, even if we grant this idea, it will be true that, within each segmented part, individuals and properties are represented holistically. Thus even granting privileged spatial decompositions, icons still do not represent an individual as having multiple properties by means of a canonical decomposition into separate constituents standing for (i) the individual and (ii) for each represented property.

The core difference between iconic and discursive formats is that the latter do, and the former do not, break down into recombinable parts each of which can stand uniquely for a particular property or individual. Any part of Figure 3 that you touch will represent a certain spatial location in the scene, and will represent various color values at that location (as well, perhaps, as shape values, higher-level properties, and individual objects). In an icon, features compose with one another and with individuals in a holistic fashion, i.e., not in virtue of combining distinct constituents (see Chapter 4, ICONIC REPRESENTATION).

2.3—Icons and complex demonstratives. Before moving forward, it is important to clarify the dialectic. The term 'iconic' has been used for decades in cognitive science to refer to representations with the properties just mentioned, as is most clearly articulated by Stephen Kosslyn and Jerry Fodor, as well as Ulric Neisser's (1967) introduction of the term to cognitive science to describe Sperling's (1960) results (see the following section for discussion). These properties play a foundational role in Kosslyn's influential theory of mental imagery (e.g., Kosslyn 1980, 34), perhaps the most important and detailed theory of iconic mental representation in cognitive science. It would thus be misleading in the extreme to use the term 'iconic' in a way that did not commit to

representations with these properties without providing an alternate theory of iconicity. It is often unclear to what extent anti-pluralists use the term in the standard way. An exception is Carey, who explicitly characterizes icons as representations whose parts correspond to parts of the scene (2009, 452) and seems to endorse the claim that icons encode properties holistically (2009, 459; cf. Block unpublished).

Burge suggests that perceptual states represent in a fashion akin to complex demonstratives, e.g., "that brown body" (Burge 2010b, 40). Since complex demonstratives have canonical decompositions—in the case of 'that brown body', into 'that' and 'brown body', and again into 'brown' and 'body'—this may suggest that Burge is committed to canonical decomposition in perceptual representation and denial of holistic binding. Complex linguistic demonstratives, however, are not iconic in any sense at all, so it would be misleading to draw that analogy in characterizing the format of perceptual representation. Partly for this reason, Burge may be best read as discussing the *contents* of perceptual representations, and not their *syntactic* properties (indeed, his use of underlines in the quotation above is meant to denote contents).

Burge's writings on perception are sometimes ambivalent as to whether he is discussing the structure of the vehicles of perceptual content or the structure of content itself. As I read Burge, this is because he in fact rejects the standard distinction between vehicles and contents.

Burge claims that vehicles cannot be individuated by their internal functional role and that (anti-individualistic) contents are constitutively 'part of the individuation conditions of perceptual states in a way that renders the vehicle-content distinction problematic (Burge 2010a, 95n45). He thus describes Fodor's (1975) hypothesis that the format of thought is discursive and quasi-

linguistic as merely a "useful idealization" (Burge 2010a, 95n45).[17] In light of this aspect of his

theory, I think Burge cannot be understood as quantifying over separate representational vehicles,

---

[17] Burge does not provide extensive argument for this position, and it seems inconsistent with cognitive science both in practice and at its foundations. Cognitive science requires a characterization of the syntactic features of mental representations in a way that abstracts away from their content to explain all sorts of mental phenomena, such as deductive inference (e.g., Braine and O'Brien 1998, Quilty-Dunn and Mandelbaum unpublished b). The fact that content is a core part of psychological explanation does not mean that we cannot discuss the internally individuated, computationally efficacious syntactic features of mental representations. One fundamental reason for needing syntax is that, as Burge would agree, contents typically supervene on relations to the distal environment. Mental representations, and indeed any computationally efficacious symbols, are causally efficacious within the mind and therefore must be operated on in virtue of their locally individuated features. Mental representations are symbols whose local properties are independent of their wide contents, and whose internal computational role is determined by those local properties. These local properties minimally include the structural features of the symbols themselves, i.e. syntactic properties. Moreover, computation is fundamentally a syntactic notion, and Burge does not provide reasons to overthrow the standard syntactic characterization of computational processes (cf. Rescorla 2012).

      Burge says we may "search textbooks and articles in perceptual psychology in vain to find mention of purely syntactical structures" (2010b, 96). But proponents of locally individuated vehicles and their syntactic features can (and almost always do) hold that those vehicles *also have contents*. It is not a prediction of the standard syntax-semantics distinction for mental representation that scientists will talk only about one in isolation of the other—and, we might note, one will also search the scientific literature in vain to find mention of purely semantic properties that are not instantiated in mental symbols.

      If, instead, Burge means to claim that perceptual psychologists never describe the syntactic features of the mental representations in their theories, then there are many significant counterexamples. To name just a few: Kosslyn's (1980) theory of mental imagery, Biederman's (1987) theory of recognition by components, Pylyshyn's (2003; 2008) theory of visual indexing, and Carey's (2009) theory of core cognition and its relation to perception are all cast in explicitly syntactic (as well as semantic) terms. Indeed, the entire imagery debate could hardly be more explicitly about the syntactic features of mental representations (see, e.g., the explicit distinction between the syntax and semantics of iconic representations in Chapter 1 of Kosslyn et al. 2006). To pick a textbook at random, Bruce et al. (2014) describe the idea of "structural descriptions" as a kind of "representational format" used in pattern and object recognition that involves mental analogues of "propositions" which are "symbolic" though not couched in natural language (2014, 272). And it is common in textbooks as well as theoretical books and articles to oppose this kind of view of object recognition to a non-discursive template-based one, a difference in syntactic structure (e.g., Ullman 1996). Outside of perception and imagery, scientific appeals to syntax in describing the structure of propositional thoughts abound (e.g., Braine and O'Brien 1998; Johnson-Laird 2006). As Burge would insist, all of these theories quantify over contents; but they clearly quantify over syntactic structures of vehicles as well.

      Aside from its inconsistency with cognitive science, ruling out any locally individuated syntactic features of mental representations creates two foundational problems for Burge's approach to perception. First, it is not clear on this purely semantic theory how contents are realized in the mind. On a broadly Fodorian picture, contents are instantiated in independently (and locally) individuated vehicles, but on Burge's picture, vehicles lack independent individuation conditions. Without content-independent vehicles that have contents, Burge's theory lacks an account of how abstract contents are realized in concrete minds. Second, it is not clear how computational processes can operate on purely semantic entities, particularly when we suppose that contents have wide conditions of individuation that are thereby inaccessible to internal computational systems. Again, the standard picture easily avoids this problem by appealing to locally individuated syntactic features. It is therefore unsurprising that, in both of these cases, Burge regards it as simply "primitive" (2010b, 96) that mental states have representational contents and that computational processes are sensitive to those contents.

one pointer-like demonstrative symbol and another attributive symbol, when making the analogy to complex linguistic demonstratives. Given his view about the relation between contents and perceptual states, I do not think he can properly be understood as discussing representational format at all—at least not in the standard, essentially syntactic sense that cognitive scientists and philosophers like Fodor (1975), Kosslyn (1980, 1994), Pylyshyn (2003, 2008), and Carey (2009) have in mind. For example, Burge (2005) claims that photographs contain singular elements, but unlabelled photographs obviously do not involve *syntactic* demonstrative items.

One might try to supplement Burge's theory with a realist, content-independent approach to syntactic properties such as representational formats and argue that perceptual vehicles are structurally similar to complex demonstratives and are also iconic. Such a project faces an uphill battle, since it would seem that to say that perceptual states involve stringing together discrete symbols into a complex, accuracy-evaluable representation just is to say that perceptual states are discursive.

One could respond by arguing that the demonstrative elements of perceptual states are non-iconic but the elements that subserve attributive contents (e.g., brown body) are iconic. But it is far from clear how a non-iconic representation can be syntactically concatenated with an iconic representation, given that they are in different formats. At first glance (and without a spelled-out account), it seems that to suppose that a representation in one format could be occurrently concatenated with another in a single, accuracy-evaluable structure threatens to be incoherent. For two representations to have different formats is for them to have syntactically divergent forms, to be in different mental languages; stringing together symbols in different languages, in the absence

of an overarching interpretation of the separate symbols, will produce gibberish.[18] Nonetheless, this version of the view will be considered in Section 5.

Perhaps a clarification of the structure of perceptual vehicles on Burge's theory or modified versions of it is forthcoming. We have good reason to be pessimistic that such a clarification can both be coherent and preserve the claim that perceptual representations are iconic.

Fortunately for dialectical purposes, Burge explicitly endorses the claim that icons and perceptual states both compose features holistically:

> Visual perceptual representation has a structure relevantly like that of pictorial representation—supplemented in various ways. Take a visual representation of an environmental edge—the border of a surface. Imagine that the representation has the form of a drawing of the edge. The drawing is a line. Such a drawing would represent not only the edge, but also its shape, length, and orientation—all as such. […] Just as one cannot draw a line without drawing its length, shape, and orientation, one cannot visually represent an environmental edge as such without representing its length, shape, and orientation, as such.
>
> (Burge 2014b, 493)

I will therefore continue to assume in what follows that parts of icons (and their distance relations) correspond to parts of scenes (and their distance relations) and that icons encode properties holistically.[19]

---

[18] One might argue that examples like the 'bed' example at the outset of this section are counterexamples. But the 'bed' case only works because the drawing functions like a word in the context of the sentence, and because the author puts the picture in that position in the sentence with the intention of indicating to her reader that he should read the referent of the picture as linguistically represented in the sentence. Such cases of format-switching require some such metalinguistic interpretation of the symbols and recognition that they are being used in an unusual fashion, and that sort of factor is not available in an account of the format of perception. If the symbols in question were *systematically* concatenated without need for such interpretation, then they would fail to be in different formats (compare emergent languages such as "Spanglish").

[19] In this chapter, I will not consider "maplike" format as something separate from iconic format (Burge and Block seem to use these terms interchangeably). I will address this issue in Chapter 3 and argue that maps are simply complexes of icons that represent a certain spatial terrain combined with discursive symbols that predicate features of locations on the represented terrain, and so provide no comfort to the view that perception is non-discursive. For

2.4—Other aspects of iconic representation. There are other claims often made about iconic representations. Some claim that icons are isomorphic, in some deep sense, to what they represent (e.g., Kulvicki 2004), or relatedly, that they resemble what they represent. There is a clear sense in which this is true: if each part of an icon represents a part of the scene and distance relations among parts of the icon track distance relations among parts of the scene, then while there may not be a perfect one-one correspondence between icon-parts and scene-parts, there is surely some meaningful sense in which the structure of the scene is mirrored in the structure of the icon. This structure-preserving relation between icons and scenes may be the intuitive fact underlying isomorphism-based and resemblance-based accounts. Goodman (1976) argues, however, there is no reason to suppose icons must resemble their scenes in some more substantive sense, e.g., that the resemblance relation is symmetric or transitive (cf. Greenberg 2013).

Shepard and Cooper (1982) describe iconic computations as exhibiting a second-order isomorphism with respect to the represented scene, since there is a one-to-one correspondence between stages in computational manipulations of icons and stages in the physical transformations described. The best example is mental rotation, which will be described in detail below; stages of mental rotation map onto stages of physical rotation so that there is a second-order isomorphism among images and what they represent.

Another claim is that icons are analog, as opposed to digital, representations (Haugeland 1998, 75–88, Dretske 1981, 135–141). The term 'analog' is used to mean different things (see

present purposes, it should be enough to note that maps satisfy Fodor's (1987) minimal criteria for the language of thought (i.e., discursive format) as Blumson (2012) argues (see also Camp 2007, Rescorla 2009).

Chapter 4, ICONIC REPRESENTATION). Sometimes it means that the representation is *continuous*, in the sense that there are no gaps between representational parts, and that this continuity corresponds to some continuity in the scene (Haugeland 1998, 82–83). Lewis (1971) argues, relatedly, that analog representations use magnitudes to represent magnitudes. This sense of 'analog' is at work in Carey's (2009, 118–135) discussion of the "analog magnitude system" for representing approximate cardinality. Analog representations in Lewis's sense need not be continuous (and so need not be analog in Haugeland's sense). Dretske uses 'analog' to describe a purported feature of perceptual representation according to which it always carries both *nested* information about abstract features (e.g., a perception of some determinate shade, e.g., red$_{241}$, "nests," and thus implicitly represents, *red* as well as simply *color*), as well as *non-nested* information (e.g., a perception cannot represent *red* without also representing some shape), while "digital" representations like thoughts carry no non-nested information except what is explicitly represented (Dretske 1981, 136ff). Kulvicki (2004), using Dretske as a jumping-off point, argues that the nesting of information in analog representations is sufficient for them to be isomorphic to what they represent.

Dretske's condition on analog representations carrying non-nested information arguably captures the fact that icons compose features holistically; if each part of an icon represents multiple properties at once, then representation of one feature will have to also represent another feature, and thus will necessarily carry non-nested information. Something very much like Dretske's analog-digital distinction will therefore be used here in distinguishing iconic and discursive formats.

Briscoe (2016; see also Hopkins 1998) argues in favor of a resemblance-based account. According to Briscoe, a two-dimensional pictorial surface determines a virtual three-dimensional "model" that resembles a three-dimensional scene in virtue of triggering similar visual experiences. This account is harder to push into the mind and apply to mental icons, as Briscoe notes (2016, 76–77). The usefulness of positing iconic representations in mental imagery (see the following section) therefore mandates formulation of a notion of iconicity that is less restrictive.

Gabriel Greenberg (2013), who rejects resemblance-based accounts, has instead argued for a projection-based account of pictorial semantics, according to which the content of a picture is the set of possible scenes (which are, for Greenberg, worlds indexed to a perspective) that can be geometrically projected from the picture in accordance with some convention of projection. If a projection-based account is correct, then perhaps it is constitutive of icons that they are projectible to some set of possible scenes according to projection conventions—though as with Briscoe's account, it's unclear whether this account could transfer to mental representations, where notions of projection and convention may not apply.

Despite their considerable interest, I won't delve into these issues here (though I will discuss them again in Chapter 4, ICONIC REPRESENTATION). The defining features of iconic representations as articulated by Kosslyn, Carey, Fodor, and others, is that there is a correspondence relation between parts of the icon and parts of the scene, and that parts of the icon encode multiple properties at once—i.e., icons compose features holistically. This minimal sense of iconicity will be at issue in what follows. Perhaps perception lacks some of the more robust properties attributed to icons in the preceding three paragraphs, and the claim that perception is not wholly iconic could

be made on those grounds. But the current project is to show that some perceptual representations are iconic in this minimal sense, and some are not.

The best way to substantiate the claim that parts of icons correspond to parts of the scene and represent multiple features holistically is simply to show its explanatory usefulness. In the following section, I will survey multiple forms of evidence that support the thesis that perceptual systems output mental representations that are iconic in this sense.

§3. Evidence for Perceptual Icons

I'll now argue that there is considerable non-demonstrative evidence in favor of the hypothesis that at least some of the representations outputted by perceptual systems are iconic (i.e., are 'perceptual icons', for short). It's important to note at the outset that the empirical evidence mentioned below in favor of perceptual icons does not necessitate that *all* perceptual representation is iconic. Even such a staunch defender of iconic format in perception and perceptual imagery as Stephen Kosslyn still affirms that "there are two sorts of representations underlying images, one 'perceptual'…and one discursive" (1980, 142), the latter constituting "a list of facts in a propositional format" (1980, 144). While the considerations below do suggest that some perceptual representations are iconic, they do not decide in favor of the anti-pluralist view that perception is wholly iconic. I will ultimately argue for perceptual pluralism, according to which perception outputs both iconic and discursive representations.

3.1—Mental imagery. Arguably the strongest evidence in favor of iconic mental representations is to be found in the experimental literature on mental imagery. The imagery debates have been

raging since the early 1970s, with Kosslyn, Shepard, and others defending the iconic interpretation, and Pylyshyn, Dennett and others defending the discursive interpretation. Phenomena like mental rotation (e.g., Shepard and Metzler 1971) and image scanning (e.g., Kosslyn et al. 1978) seem to show that differences in processing speed are proportional to spatial differences in what is represented, suggesting that the processing involves manipulating or "scanning" iconic mental images with functional analogues of the spatial properties of the represented scene.

In Finke and Pinker 1982, for example, subjects were shown a picture containing four dots; then the picture disappeared and was replaced by an arrow pointing in a specific direction, and subjects were asked whether the arrow was pointing at one of the previously presented dots. Reaction times increased proportionally to the degree of distance between the arrow and the dot. Furthermore, though subjects were not instructed to form mental images, when asked afterward how they completed the task, all but one of the twelve subjects reported using mental imagery. The iconic interpretation of this effect is that the subjects held an iconic perceptual representation in memory and imaginatively superimposed the arrow on it. The proportionality between distance and reaction time is explained by appeal to the functional–spatial structure of the icon itself.

This explanation invokes iconic representations with parts that correspond to parts of the scene and that encode properties holistically. In order for subjects to systematically take longer when reporting on greater distances between the arrow and the dot, subjects must need to access intermediate location values one after another such that a greater number of intermediate location values takes more time. There must therefore be parts of the representation that correspond to

locations in the scene such that scanning over a longer distance requires scanning over more parts of the image.

Moreover, parts of the representation not only correspond to parts of the scene, they also encode both the location and any shape present at that location. The functional relationship between the part of the representation that represents the arrow and the parts that represent each dot is explained by the fact that each of those parts also represents a certain location value, thus situating representation of shape within a larger functional–spatial array. Iconic explanations of image-scanning therefore invoke the holistic character of icons.

The same is true of iconic explanations of mental rotation. In mental rotation cases, subjects see a pair of objects and have to determine whether they are different objects or the same object but merely rotated at different orientations. Successfully identifying a match takes longer if the degree of rotation is greater. This suggests that there is an operation of mental rotation that is functionally analogous to physically rotating an object, where transitioning from a representation of an object at rotation angle A_m to representing it at angle A_n requires representing it at all or many of the angles between A_m and A_n.

On an iconic explanation the representations that are mentally rotated have a functional organization that recapitulates the spatial structure of the imaged objects, and must therefore have parts that correspond to spatial parts of those objects.

Furthermore, parts of the image that represent parts of the object must also represent properties like shape, size, and orientation that are used to detect whether the object is the same. The shape of the object cannot be accessed in a way that is not bound up with the orientation and

relative locations of parts of the object; if it could, then the shape could simply be recovered and matched to the other object without taking extra time corresponding to the degree of rotation. Iconic explanations of mental rotation therefore invoke the holistic character of icons as well.

The mental imagery literature thus not only provides good evidence in favor of iconic representations, it also vindicates the notion of iconic representation used in this dissertation.

Pylyshyn (2002; 2003) has objected strenuously to iconic interpretations of mental imagery, though even he admits that Kosslyn's iconic model "can be taken as the received view in much of the field" (2002, 181). Pylyshyn often explains effects like these in terms of experimenter demand (i.e., subjects guessing that experimenters want them to take more time for longer distances, for example) or tacit propositional knowledge of what objects look like. I will not engage at length with his arguments here. However, like other discursive models of mental imagery, his explanations of these experiments are *post hoc*—first the effect is observed, then a proposed explanation is that subjects might have, e.g., simply drawn on tacit knowledge or guessed what experimenters were up to. But this sort of *post hoc* explanation is always possible in a scientific field in which the evidence dramatically underdetermines the underlying mechanisms, a methodological conundrum that can be observed in nearly every nook and cranny of cognitive science. It is important in these cases that proposed explanations cannot merely be cooked up in response to the evidence, but also that they predict new evidence in systematic ways. Iconic models of mental imagery drove results like Kosslyn's and Finke and Pinker's in the 1970s and 1980s, while alternative explanations in terms of discursive format would arrive after new evidence had been

presented. The fact that iconic models led the experimental research by making predictions that were consistently vindicated points in favor of those models over *post hoc* alternatives.

The argument presently on offer is that there is good reason to think that visual images are iconic, and that if there is good reason to think visual images are iconic, then there is good reason to think that at least some representations in visual perception are iconic. This is perhaps a peculiar way of arguing, since it is often just assumed by proponents of the iconicity of mental imagery that representations in imagery are of a piece with representations in early perceptual processing, which are in turn assumed to be iconic. According to Block, this is assumed widely enough that the putative similarity of perception and imagery is often erroneously taken to be just what's at issue in the imagery debates (1981, 9). Though it isn't *per se* what's at issue, some defenders of the iconicity of imagery build the recruitment of perceptual representations into their account of imagery as iconic. Kosslyn et al., for instance, stipulate that their use of the term 'mental image' is such that "a mental image occurs when a representation of the type created during the initial phases of perception is present but the stimulus is not actually being perceived; such representations preserve the perceptible properties of the stimulus and ultimately give rise to the subjective experience of perception" (2006, 4). Kosslyn is at times even stronger, writing, "All characterizations of imagery rest on its resemblance to perception" (1995, 267). It is thus built into prominent models of visual imagery that representations in the same format are delivered by early perceptual processes.

The relatedness between perception and imagery is often cast in two ways: first, that imagery is functionally and neurally related to perception, and second, that conscious imagery is

phenomenologically related to conscious perception. Assuming mental imagery is iconic, we can identify two separate but connected arguments for the iconicity of (at least some) perception.

Recruitment argument:

(1) Mental imagery is iconic.

(2) Mental imagery recruits representations from early perceptual systems.

(C) Representations in early perceptual systems are probably iconic.

Phenomenological argument:

(1') Mental imagery is iconic.

(2') The phenomenology of mental imagery reflects its iconicity.

(3') Representations in imagery have the same type of phenomenology as some representations in perception (i.e., what it's like to visualize a candy cane and what it's like to see a candy cane are phenomenally similar).

(C') Some representations in perception are probably iconic.

Premises 1 and 1' were addressed earlier and will be assumed for the sake of argument. Premise 2 in the recruitment argument is an empirical hypothesis, based in part on neuroscientific evidence about which cortical areas are used in perception and imagery (Kosslyn et al. 2006, chapter 5). One simple justification is well known, viz., that vision and visual imagery compete for resources, in both cortical and psychofunctional terms (Pearson et al. 2015). When trying to visualize something, people often close their eyes or look up, effectively reducing salient input. One classic explanation of this is that the same psychological systems that underlie visual perception also underlie visual imagery, so loading visual perception with salient input (i.e., keeping your eyes

open and looking straight ahead) will suck up resources and thus hinder the production of visual mental images; once the eyes are closed or looking away from salient stimuli, visual resources are freed up to be used for imagery. If this simple explanation is correct, then since the same systems (viz., early visual systems) are used for vision and visual imagery, the types of representations used are probably the same. So, if visual images are iconic and visual imagery uses the same systems as early vision, then the outputs of early vision are probably iconic.

Another example of the commonality between perceptual representations and mental images involves perceptual learning. Subjects who merely imagine relevant stimulus properties become better at discriminating those properties. Grzeczkowski et al. (2015) presented subjects with three vertical lines, and asked them to indicate whether the line in the middle was offset to the right or left (see Fig. 4). In one condition, the subjects trained by seeing a series of stimuli in which the middle line was not offset at all, and they were instructed to imagine the middle line as offset to the right upon hearing a 1000hz tone and to imagine it as offset to the left upon hearing a 700hz tone. These subjects showed enhanced discrimination after this training despite the fact that it was merely imagined (see also Tartaglia et al. 2009). This experiment suggests a deep commonality between perceptual and imagistic representations, which is predicted and explained by the hypothesis that perceptual and imagistic representations are both couched in the same iconic format.

Figure 4—Stimulus used by Grzeczkowski et al. (2015). Subjects were instructed to imagine the

middle line as offset to the right upon hearing a 1000hz tone, and to the left upon hearing a 700hz tone. Afterwards, they showed improved performance on a task that involved judging whether the middle line was offset to the right or left.

The phenomenological argument is trickier. Premise 3' of the phenomenological argument seems right, even undeniable, to many people. We often use perceptual terms to talk about imagery: we say we can "see her face now" when talking about visualizing the face of a relative, or that we can "taste it now" when talking about recalling a pleasant (or unpleasant) taste. Subjects tend to describe what they're doing in classic imagery tasks in perceptual terms, even when otherwise unprompted to do so (e.g., Finke and Pinker 1982, 144). The obviousness of the perceptual character of the phenomenology of imagery has driven some theorists to characterize the imagery debate as just a debate about whether the perceptual phenomenology experienced in imagery tasks is "epiphenomenal" (Block 1981, 6–9; though Block points out the inaccuracy of this characterization).

What is less clear is whether this commonality in phenomenology requires a commonality in representational format—i.e., whether premise 2' is true. One reason to think it is true is that the phenomenology in imagery tasks seems to track manipulations of the image in respect of its iconic features. Take a classic informal demonstration of mental imagery: How many windows were there in your childhood home? To many people, it doesn't seem like they simply recall the fact that there were a certain number; rather, according to the iconic theorists, you call up a mental image of your home, and then "scan" it like an image. What's more, it seems phenomenologically

like that's exactly what you're doing.[20] And while, according to the theory, you scan from one window to the next, the phenomenology changes accordingly. We need not suppose we have perfect privileged access to our mental imagery to recognize this (cf. Schwitzgebel 2002). It seems, subjectively, like the phenomenology tracks the iconicity of the representations in just the same ways that it shows up in behavior. So the key premises in the phenomenological argument seem to have some independent support.

Many people report no perception-like imagery at all (Phillips 2014). This fact might be taken to undermine premises 2' and 3'. However, the phenomenological argument does not claim that phenomenology is responsible for iconicity or drives the relevant behavioral effects. Instead, the claim is that the phenomenology of imagery in typical subjects with conscious imagery both reflects its iconicity and is subjectively similar to perception, and that this face provides evidence that perception is iconic. Reports of no conscious perceptual imagery are rare, and people who report less vivid mental imagery still report perception-like imagery while also exhibiting the same behavioral effects on imagery-based tasks (McKelvie 1995). Furthermore, it is unclear whether non-conscious-imagers experience a distinct, non-iconic phenomenology as opposed to simply experiencing no imagery-based phenomenology at all. Only the former possibility, however, would seem to threaten the phenomenological argument (and even then, only for non-conscious-imagers), since it is compatible with the claim that the phenomenology of imagery, when it is present, reflects the iconicity of imagery as well as its similarity to perception.

---

[20] There is the wrinkle here that some people do not report having mental imagery. But the focus here is on those who do, and the phenomenal character of their mental imagery.

Before moving on, it is worth addressing a puzzling aspect of the empirical evidence on these issues. One might think it's odd to infer from the iconicity of imagery to the iconicity of perception since, as mentioned above, imagery theorists often cast the iconicity view in terms that assume early perceptual representation is iconic. That is, an argument for the iconicity of imagery implicit in some theories is that *perception* is iconic and imagery is like perception so, therefore, *imagery* is iconic. But the direct experimental evidence is arguably stronger and greater in volume for the iconicity of imagery than for the iconicity of perception. Why did the evidence turn out this way? One possible explanation is that representations in perception are not open to free manipulation by the subject in line with task demands—for perception, either the right stimulation occurs or it doesn't. Since representations in mental imagery can be triggered in a relatively controlled way by subjects in the absence of stimuli, it is experimentally easier to hold the representation fixed and manipulate it in various ways to make its format surface in behavior. When you're actually looking at your childhood home, you don't scan a mental representation of it; you scan *it*.[21] Thus it might not be so surprising that it turned out to be easier to directly behaviorally show iconicity in mental imagery than in perception.

3.2—Iconic memory. One variety of short-term memory in vision is called "iconic memory" (Neisser 1967). The original motivation for positing iconic memory was the famous Sperling (1960) experiments, in which subjects are presented with (for instance) three rows containing three

---

[21] The use of 'scan' here is loose—defenders of the iconicity of mental imagery need not posit a homunculus that looks at the mental image as though it were a literal picture. And when you visualize the face of a relative, the thing you're most directly conscious of is typically the face, not the mental representation. But in the case of imagery tasks such as counting the windows, the representation is often an object of attention, whereas ordinary perception typically seems "transparent." That's arguably why counting the windows in imagery seems more like scanning a picture whereas counting the windows while looking at the house seems more like scanning the house itself.

letters each (see Fig. 5). When asked to report all nine letters, subjects can only provide a few of them. But when a row is cued after the brief presentation is gone (by an arrow pointing toward the middle row, for example, or a low tone indicating the bottom row), subjects can get all or nearly all the letters in that row, no matter which row is cued. There must be some stored representation that preserves most or all of the image and allows subjects to "read" letters off it despite outstripping what the subject can explicitly individuate at a time. This effect has often been taken to implicate an iconic visual representation stored in some visual memory system that, when cued, is available for conceptualization (e.g., Neisser 1967; Fodor 2007).
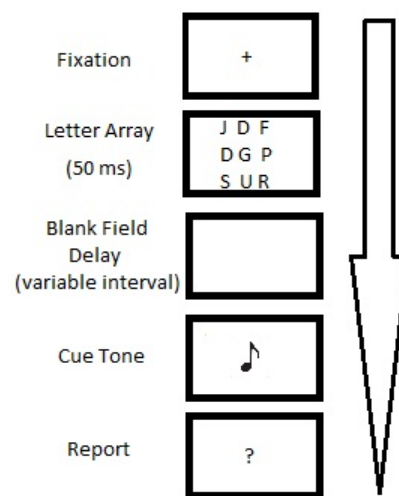
Figure 5, Sperling's (1960) partial report paradigm[22]

Since icons don't break down into separate symbols for each represented item, then an icon that represents a scene brings along for free an array of properties represented in that scene. Since parts of the icon correspond to parts of the scene and represent multiple properties simultaneously, there aren't distinct symbols for each individual represented item separate from the holistic

---

[22] Taken from Wikimedia Commons.

binding of properties (including location). Introducing a greater number of items into the scene thus doesn't require adding symbols; it just requires shifting around the properties that are represented at various locations in response to the properties of the new items. For this reason, icons can encode a large amount of information without greater cost. Thus the holistic character of icons can provide an explanation of iconic memory.

Just as there is an iconic short-term memory system in vision (Sperling 1960; Coltheart 1980; Irwin and Yeomans 1986; Becker et al. 2000; Bradley & Pearson 2012; Persuh et al. 2012), there are analogous iconic short-term memory systems in other modalities: "echoic" memory in audition (Neisser 1967; Darwin et al. 1972), tactile memory in tactition (Bliss et al. 1966), and possibly short-term olfactory memory (White 1998; cf. Stevenson & Boakes 2003). There are important differences between these systems, especially temporal capacity—e.g., visual iconic memory lasts for hundreds of milliseconds, and auditory echoic memory lasts for several seconds. But they are arguably alike in that they involve the storage of unconceptualized iconic representations.

It is possible that these effects could be explained by massive parallel processing that delivers discursive representations into a very large memory store. This explanation is inelegant in that there is no independent reason to think that discursive representations can be stored in such high capacity, particularly given the fact that there is a very small item limit on discursive representations in visual and verbal working memory systems (e.g., Cowan 2001; note that this point does not require endorsing a "slot-based" over a "resource-based" model of working memory—see Suchow et al. 2014). Moreover, it leaves unexplained why there should be such a

sharp difference between the storage capacities of iconic memory and later visual working memory, given the putative commonality in representational format.

There is also independent reason to think that iconic memory stores perceptual icons. The sensory "informational persistence" (Coltheart 1980) that underlies iconic memory utilizes early, retinotopically mapped visual cortical areas of the brain (Duysens et al. 1985; Irwin & Thomas 2008). These areas of the brain are also the loci of visual images (Pearson & Kosslyn 2015). This suggests that the format of representations stored in iconic memory is the same as the format of representations used in mental imagery, in which case we should conclude that iconic memory stores perceptual icons.

Phillips (2011) aims to refute this last point. He argues against the idea that the format of representations stored in iconic memory are genuinely iconic by appeal to differences between iconic memory and visual imagery. He argues first that "subjects in Sperling's task are not instructed to make, and do not consistently report making, use of visual imagery" (2011, 211). This argument is far from decisive given that the use of imagery does not entail reportability of such use (McKelvie 1995). Moreover, the storage and access of an iconic representation in online perception is not the same thing as the top-down formation of a mental image. Positing the same representational format in both cases does not require that the mechanisms by which the representations are formed, stored, and accessed are the same. The fact that these different processes utilize representations in the same format does not entail that subjects that can report on one process can report on the other, nor does it entail that hallmarks of one process should be exhibited by the other.

This point undermines Phillips' second and *prima facie* more substantive argument. This second argument relies on a study by Brockmole et al. (2002). Brockmole et al. presented subjects with a grid of squares wherein nearly half of the squares were filled in with dots; after a blank screen, subjects were then presented with another grid with different squares filled in, such that there is a single square that is not filled with a dot in either the first or second grid. The task is to identify this square. Participants are good at the task when there is no gap in time between the two grids (i.e., they are accurate on 79% of trials), but there is a significant decrease in accuracy with a 100ms gap (21% accuracy). At longer durations, however, accuracy jumped back up (roughly between 60% and 70% accuracy). Moreover, the jump in accuracy stabilizes after 1.5 seconds, which corresponds to the amount of time it takes to generate a mental image (Kosslyn 1980). Phillips argues that, if iconic representations were stored in iconic memory, than we should expect a similar time-course for iconic memory; however, iconic memory exhibits the opposite time-course, since partial-report superiority obtains well under a second and fails above a second. Thus, he concludes, "the imagistic account is not viable" (Phillips 2011, 212).

This argument assumes that, on iconic models of iconic memory, the mechanisms underlying the generation of an icon in online visual perception and those underlying the generation of a visual image must be the same. This assumption is unwarranted. The generation of a visual image is typically a voluntary, top-down process. According to standard iconic models of mental imagery (e.g., Kosslyn 1980; 1994; Kosslyn et al. 2006), the generation of a mental image requires the prior access of discursive information that is used to voluntarily construct a mental image which can then be stored, manipulated, and accessed. This takes time. The formation of an

icon in online visual perception, however, is not top-down and need not access stored discursive information. Iconic explanations of Sperling-type effects thus do not make the predictions about time-course that Phillips ascribes to them.

One might still wonder why performance on the Brockmole et al. experiment improves after a second while performance on the Sperling experiment has tanked by that point. The view defended here indeed requires that both experiments involve the use of iconic representations. Nonetheless the tasks impose quite different demands. The icon in the Sperling case just needs to be attended such that a subset of its information can be transformed into reportable conceptual states. In the Brockmole et al. case iconic memory has been masked and thus wiped clean by the presentation of the second grid (which explains the drop in accuracy at 100ms). Subjects must then engage in top-down generation of a mental image, which takes over a second. Indeed, Lewis et al. (2011) replicated Brockmole et al.'s finding while explicitly instructing subjects to form mental images. This effect strongly suggests that the time-course of the Brockmole et al. task is due to the top-down formation of a mental image, and is therefore irrelevant to the question of what formats are deployed in online visual perception and stored in iconic memory.

This line of response to Phillips highlights one of the core architectural principles of this dissertation: Whether two processes are type-distinct and whether two representations are type-distinct are separate questions. A concept can be deployed in cognition and, I'll argue, can be deployed in perception. Nonetheless, perception and cognition run distinct types of processes. Similarly, an iconic representation can be deployed under top-down cognitive control in mental imagery and take over a second to generate, or it can be deployed in online visual perception and

be generated within 50ms. The format of a representation cannot necessarily be read off the process in which it is deployed.

It is worth reiterating that these arguments for perceptual icons are not intended to be decisive or deductively valid. The argument is simply that these effects are best explained by positing perceptual icons. Laurence and Margolis (2012) argue that it is a "conceptualization fallacy" (304) to infer from the fact that (e.g.) the letters have to be conceptualized to get their identities that they were not already conceptualized (as shapes, perhaps) prior to the conceptualization of their identities. But the argument is not that it is *impossible* that the letters are already conceptualized; that would indeed be a fallacious inference. The argument is that the stored representations exhibit a very high capacity and persist when not conceptualized for report, and that these properties are best explained by positing an iconic format. And moreover, the mere possibility that the letters are conceptualized in iconic memory is uninteresting, given the lack of positive evidence for that conjecture. The appeal to perceptual icons might not turn out to be the best explanation after all, but Laurence and Margolis only raise the in-principle possibility that the explanation is false, which does not suffice to render such abductive arguments fallacious.

3.3—Ensemble perception. People can only explicitly individuate about four individual objects at once (Pylyshyn 2003; object perception will be discussed at length below). Nonetheless, there is evidence that we extract statistical regularities in scenes by computing over many more than four objects. This capacity is known as ensemble perception.

In a pioneering study, Ariely (2001) showed subjects as many as 16 circles of varying diameters. After the circles disappeared, subjects were shown a probe circle and asked whether it

was identical to one of the 16 just seen. Subjects were more likely to answer "yes" the closer the diameter of the probe was to the mean diameter of the set (independently of whether the probe actually was a member of the set). This result requires subjects to implicitly encode the average diameter of the set of circles even though they do not seem to encode each individual circle as such. Furthermore, subjects were also above chance when asked explicitly whether a probe was greater or smaller in diameter than the average of the set.

Similarly to iconic memory, ensemble perception effects like this suggests that subjects represent an array of items without deploying a discursive representation for each individual item that can be stored and used for report. If subjects deployed discursive representations for each individual item, then they should succeed at identifying an object as a member of the set. But Ariely found that "observers were unable to distinguish test spots that were in the set from those that were not" (2001, 159). Remarkably, this was found even when the set contained only four circles.

This sort of ensemble coding has also been found for color, location, and even facial emotions as well as many other properties (see Haberman & Whitney 2012 for an overview).

Ensemble perception can be explained by supposing that perceptual icons function as inputs to a process that extracts statistical averages. This explanation in terms of iconic inputs to ensemble coding can explain why subjects lack information about whether an individual item was a member of the presented set—the items were encoded iconically and were not segmented out and represented by distinct discursive representations (see also Fodor's [2007] discussion of "item effects").

3.4—Spatial representation. Another family of arguments for the iconicity of visual representation concerns the way in which spatial properties are visually represented. Burge has recently argued for the iconicity of vision by appeal to spatial perception. "Visual perception," he writes, "occurs in an ego-centrically anchored, spatial coordinate system" (Burge 2014b, 488).

> The format of visual representation takes on some of the geometry of the mapping of light on the retina. Representation via ego-centrically anchored spatial coordinate systems owes much to the spatial layout of light registration by retinal receptors. Some of this layout of pre-perceptual registration of retinal information is preserved and co-opted by the perception-formation process. Of course, both the pre-perceptual sensory layout and perceptual content with 3-dimensional significance derive from mapping the projected spatial structure of the environment. The sensory layout is given 3-dimensional spatial content in the format of an ego-centrically anchored spatial coordinate system, at stages of computation beyond retinal registration. This format informs nearly all perceptual representation.
>
> [Burge 2014b, 494]

The argument is somewhat obscure. It seems that Burge takes the particular way in which the visual system processes spatial information to involve a kind of "mapping," which results in a visual representation whose parts represent spatial parts of the represented scene. Here is a reconstruction of the argument:

Burge's spatial argument:

(1*) The geometrical structure of retinal activation registers spatial information via an iconic format.

(2*) The visual system constructs representations of the scene that preserve the geometrical structure of retinal activation.

(C) The representations constructed by the visual system are iconic.

Premise 1* can be assumed—characterization of retinal activation as an "image" is ubiquitous in psychology and neuroscience. It is less clear what grounds 2*. According to Burge, visual perception represents by way of egocentric 3D spatial coordinates. The format of these perceptual spatial coordinates is determined in some significant way by the geometrical structure of the retina. Thus, since the structure of retinal activation is iconic and visual spatial representation "preserves and co-opts" the structure of retinal activation, visual spatial representation is iconic.

The idea seems to be that there is a mapping from the geometrical properties of the scene to the retina via reflection of light, resulting in an iconic retinal image, and there is a second, similar mapping from the geometrical properties of the retinal image to a spatial coordinate system deeper in the visual system. This spatial coordinate system could, for instance, instantiate the "scenario content" described by Peacocke (1992). It seems that what is doing the work in showing visual spatial perception to be iconic is not the relation to the retina, but rather the assumption that such perception is carried out via such a coordinate system. After all, one could imagine a system that goes straight from retinal activation to discursive conceptual representation of each portion of the visual field, such that the resulting representation "owes much to the spatial layout of light registration by retinal receptors" but is not iconic—indeed, this is presumably how conceptualist anti-pluralists (e.g., Mandik 2012) think space is perceptually represented. So there must be some independent reason to think that perception represents via syntactic analogues of spatial coordinates.

One independent reason concerns the structural features of early perceptual cortical areas, to which I turn now.

3.5—Topographical mapping. The idea that early perceptual representations represent via syntactic analogues of spatial coordinates is consistent with a neural phenomenon detailed in any undergraduate vision science textbook, viz., topographical mapping. 'Topographical mapping' or 'retinotopic mapping' are terms used to describe the fact that the geometrical structure of the retina and hence of the scene is recapitulated in the geometrical structure of the primary visual cortex (see Fig. 6). Frisby and Stone, for instance, write that "adjacent regions of cortex deal with adjacent regions of the scene" (2010, 7). This cortical topographical mapping "preserves the neighborhood relationships that exist between cells in the retina" (*Ibid.*).
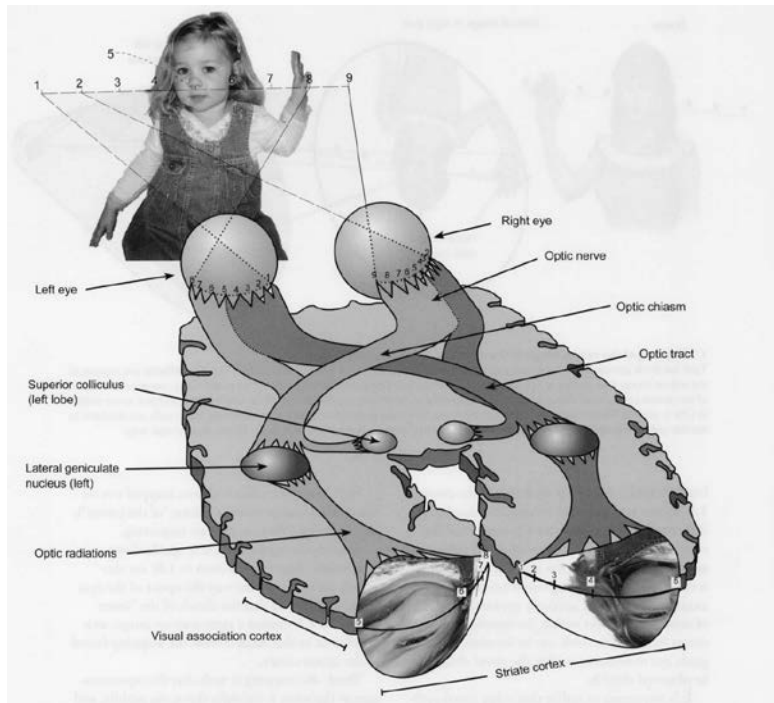
Figure 6, Retinotopic mapping in V1[23]

The retinotopic topographical mapping in V1 is of particular interest here because it seems

to be a case where the iconic format of perception emerges directly from neural structures. V1

looks like it subserves perceptual icons. Kosslyn writes that "neurons in the cortical area are

organized to preserve the structure (roughly) of the retina. These areas represent information

depictively in the most literal sense; there is no need to talk about an abstract functional space akin

to that defined by an array in a computer" (1994, 13). Finding iconicity in neural structure does

not entail that the structure subserves iconic mental representation (Pylyshyn 2003, 394ff), but it

might provide some evidence. Indeed, given that columns of cells function to detect edges at

various orientations and are organized into hypercolumns that correspond to particular locations

on the retina, it is plausible that V1 subserves iconic representations that holistically bind edges

[23] From Frisby and Stone 2010, 5, Fig. 1.4.

and contours to perceived locations (and similarly for other early visual cortical areas that represent other features via retinotopic mapping).

Though topographical mapping seems to bolster Burge's claim about the retinotopic character of visual perception, the claim is not a neurophysiological one, but a psychofunctional one. As Kosslyn put it, "there need be no actual picture in the brain to have a depiction: all that is needed is a *functional space* in which distance can be defined vis-a-vis the interpretation of the information" (1995, 282). If some visual perception represents via a spatial coordinate system, and that coordinate system is determined by a geometrical projection from the retinal image (and so is indirectly determined by a geometrical projection from the represented scene), then that gives us good psychofunctional reason to suppose that at least some spatial visual representation is iconic. The fact that iconic memory (Irwin & Thomas 2008) and mental imagery (Pearson & Kosslyn 2015) seem to be realized in retinotopically mapped visual cortex provides independent verification for this claim. Thus neuroscience, perceptual psychology, and philosophy of perception conspire in suggesting the iconicity of early visual spatial representation.

Clark (2009) argues against an iconic interpretation of topographical mapping. One of his key points is that V1 (for example) does not literally satisfy Kosslyn's claim that distance relations between parts of early visual representations map onto distance relations in the scene: "If you move half a millimeter in one direction, you might not change the receptive field at all, but instead move to a region receiving input from that same receptive field, but from the other eye." (2009, 294). But the claim is not that the spatial structure of V1 perfectly satisfies the conditions on iconicity; one could nearly deduce *a priori* that that radical hypothesis is false. The claim is that the spatial

structure of V1 enables realization of a functionally individuated mental representation that satisfies those conditions. The fact that columns of cells correspond to edge-orientations at particular locations, and that there are distance relations among columns that track distance relations on the retina, allows V1 to be exploited at the functional level in subserving perceptual icons. Asserting that V1 instantiates icons does not require that every aspect of the spatial structure of V1 reflect the icons it instantiates. This point is merely an instance of the autonomy of psychofunctional explanation from neuroscience, and does not undermine the evidential status of topographical mapping for perceptual icons.

3.6—Richness and fineness of grain. Philosophers often talk of perceptual content as being richer or more fine-grained in information than what is conceptualized at any given time (e.g., Dretske 1981, Evans 1982, Peacocke 1992, 2001, Tye 1995, 2006). One gets this impression easily in a busy environment, such as a city street. The intuition these philosophers exploit is that, in such an environment, there are many more properties visually represented than we can carve up conceptually in a moment, and in much finer detail. As mentioned in the discussion of iconic memory, since iconic representations need not segment the perceptual field and construct individual representations for each represented item, they easily convey a large amount of information to consuming processes, including quite specific information about, e.g., colors and shapes. Moreover, the representational apparatus exploited by icons need not be constrained by the discursive concepts deployable in a given creature's mind, in which case perceptual icons might be much richer or more fine-grained than what the creature can conceptualize. If it's true that

perception does exhibit richness and fineness of content, this aspect of perception would be explained by positing perceptual icons.

3.7—Perception–cognition border and top-down effects. Larger theoretical concerns might also drive one to assume perception is iconic. For one thing, if perception is iconic and cognition is discursive, those facts may suffice for there being a natural division between perceptual systems and cognitive systems, even if there is systematic cross talk between perception and cognition. Block (ms), for instance, argues both that cognitive penetration is rampant and that there is a joint in nature between perception and cognition. For Block, this joint in nature is largely due to the fact that perception is iconic and cognition is discursive.

Furthermore, if you are committed to the existence of a rigid functional border between perception and cognition (and thus deny that perception can access information stored in cognition), then you will likely be inclined to posit differences in the kinds of processing that occur in those two kinds of systems. If System A is encapsulated from System B, then it is possible that A and B run exactly the same types of processes and differ only in that A cannot access B's information. Nonetheless, it is implausible that two systems that are encapsulated from one another would be otherwise exactly alike. Thus if one assumes that perception is encapsulated from cognition then one should increase one's credence in other differences, especially differences in the types of processes implemented in the two systems. Moreover, differences in processes implemented in the two systems could partially explain why a system that implements process P cannot use information stored in a system that is unable to implement P. And differences in format can explain differences in processing.

If two systems process representations in different formats, then there will be differences in the nature of the processing itself. Computational processes operate on the syntactic structures of the representations that figure in them, and different formats are different types of syntactic structures. The format of representations operated on in a given system will therefore determine to some significant extent how processing in that system proceeds, and thus how it can be affected by other systems. If the format of perceptual representation is iconic and the format of cognitive representation is discursive, then there will be significant differences in the kinds of processing that occur in perception and cognition.

Furthermore, this difference in format would make interaction in processing across the two systems harder to achieve (though not impossible). If perception and cognition differ in format, and so the computations performed by the two systems differ in fundamental ways, we should be *a priori* disinclined to expect effects of cognition on perception. For example, if processes in perception exclusively operate on icons, then a discursive cognitive representation cannot be accessed by a perceptual process. Top-down effects of discursive cognitive representations on perceptual processing would thus require some mediating mechanism that takes discursive cognitive representations as inputs and either (a) produces perceptual icons as outputs (which are then accessed directly in perceptual processing), (b) alters the contents of already existing perceptual icons (either outputs or intramodular states) without otherwise figuring in online perceptual processing, or (c) alters the rules according to which icons are processed in perception, again without otherwise figuring in perceptual processing. If representations in perception and cognition are in the same format, on the other hand, then representations in cognition could

simply figure directly in perceptual processing without needing a mediating mechanism to perform any of (a)–(c).

To understand the relation between the format of representations in a given system and how those representations are processed in that system, it will be useful to dwell momentarily on how processing rules are built into psychological systems.[24] In central cognition, there are arguably inferential rules built into the architecture pertaining to the constituent structure of propositional thoughts. For example, thinking APPLES ARE RED and IF X IS RED THEN X IS COLORED will lead, automatically and without the mediation of any mechanism at the intentional psychological level, to the thought APPLES ARE COLORED. A modus ponens rule, *If P and IF P THEN Q then Q*, is plausibly *built into the architecture* of central cognition: that is, central cognition is organized such that when representations of a certain type (e.g., having a certain constituent structure) are tokened in the system, then, all else equal, representations of another certain type will be tokened in the system. The *ceteris paribus* clause abstracts away from performance errors and other external interventions on the system. The modus ponens rule in central cognition is an *inferential* rule in that it pertains solely to the constituent structure of propositional thoughts (Quilty-Dunn and Mandelbaum unpublished b).

Since icons lack constituent structure, transitions between icons in perceptual systems cannot be governed by inferential rules. Assuming that such transitions are nonetheless rule-governed, there must be noninferential rules built into the architecture of these systems, such as rules that pertain to the contents of the icons (e.g., for some 2D shapes a* and b* and 3D shape c*

---

[24] I use asterisks to denote iconic contents.

a rule of the form, *If a* and b*, then c**, might be used in stereoscopic depth computations). These implicit processing rules specify features of iconic representations. Discursive representations cannot, therefore, satisfy their antecedents. That might be too quick—it is possible that perceptual systems are constructed with disjunctive rules that take either iconic or discursive representations as inputs, of the form *If D or e*, then f**. In that case, either tokening the concept D or forming the iconic perceptual representation e* would cause the perceptual system to transition to f*. But in the absence of an independent reason to suppose such rules are built into the perceptual architecture, this stipulation seems unmotivated.

One kind of route for concepts to alter perceptual processing would be via some mechanism that takes concepts as inputs and delivers icons as outputs, such as in possibility (a) above. Macpherson (2012) makes a proposal along these lines, arguing that the mediating mechanism is perceptual imagination. Thinking CAUCASIAN upon seeing an apparently Caucasian face might cause the formation of an iconic visual image of a lighter color, which is then accessed by perceptual processing, yielding a perceptual representation of a face that's lighter than the actual stimulus (as is argued to be the case by Levin and Banaji 2006—see Firestone and Scholl 2015 for a rebuttal). Alternatively, the discursive representation could modulate the icons already operated on in the perceptual system without creating a new icon, in line with possibility (b).

Another kind of route would be not to alter the representations being processed, but to alter the processing rules themselves, as in (c) above. For example, believing that people with narrow nose bridges tend to have lighter skin might alter the visual system such that when icons with a certain configuration corresponding to narrower noses are tokened, then next icons tokened

will tend to represent lighter colors, and this transition is not mediated by any token conceptual states. In this case, for some shape g* and color h*, background knowledge has built rules of the form, *If g*, then h*,* into the architecture of the visual system. There must be some mechanism in this case that maps discursive thoughts to perceptual processing rules. It's unclear what this mechanism might be—though perhaps, if certain radical Bayesian theories are true (Clark 2013; Hohwy 2013; Lupyan 2015), priors stored in cognition could determine processing in perception in this sort of way.

The main moral of this discussion has been that a difference in format between perceptual and cognitive systems makes it the case that effects on the former by the latter should be expected to involve an additional mechanism of some sort, whereas a commonality of format doesn't suggest that top-down effects should be mediated by such a mechanism. The difference in format therefore supplies a surmountable but nontrivial barrier to effects across psychological systems. Settling this issue is, of course, an empirical matter. It might turn out that there *is* some mechanism that accomplishes any or all of (a)–(c); and it might turn out that perception and cognition are alike in format but perception is nonetheless totally encapsulated. The above discussion simply suggests that theorists who accept a rigid perception–cognition border and deny significant "top-down" effects of cognition on perception will, all else equal, be sympathetic to the claim that perception is wholly iconic while cognition is discursive.[25]

---

[25] Strictly speaking, the anti-pluralist iconophile is only committed to saying that the *outputs* of perception are iconic; the representations accessed in delivering those outputs may in principle be in any format. In that case, there are no restrictions on the way in which discursive representations can be accessed by perceptual processing. It would be a strange view, however, that allowed the involvement of personal-level discursive conceptual states in subpersonal processing but not the personal-level outputs of that processing.

On a perceptual pluralist view, if all perceptual processes are encapsulated, then encapsulation cannot be due entirely to perceptual icons. Nonetheless, perhaps the encapsulation of the early perceptual processes that output perceptual icons can be explained in terms of representational format. This explanation will be especially plausible if early perceptual systems are encapsulated to a greater degree than later ones, or if they are diachronically as well as synchronically encapsulated (whereas later systems might be synchronically but not diachronically encapsulated).

3.8—Summing up. Taking all of these considerations together, there is good (tentative, abductive) reason to suppose that there are perceptual icons. But as we will see below, not all representations delivered by perceptual systems are iconic.

§4. Perceptual Object Representations

There are several good candidates for discursive format in perception, some of which will be discussed toward the end of this chapter. But the bulk of my case will rest on object perception.

4.1—Object perception. Object recognition in high-level vision—e.g., recognizing something as a duck, or as your cousin, or as a cup—seems to be a species of cognition, and so involves the deployment of concepts. There is a more basic form of object perception, however, which occurs in mid-level visual processing. Kahneman, Treisman, and Gibbs (1992) showed the existence of an "object-specific preview benefit" (OSPB). In this object-reviewing paradigm, subjects are presented with two objects (e.g., circles) in which features are previewed (e.g., letters—see Fig. 7). The features then disappear and the objects move in different directions. Then a feature flashes in one of the

objects and subjects answer whether it is one of the previewed features or not.[26] It doesn't matter for the task which object the feature was originally previewed in. If, however, it is the same feature that was previewed in that particular object originally, subjects are quicker at identifying a match.

For example, in Figure 7, subjects would be quicker to recognize the 'T' as a match if it were presented in the same object in both the preview and target displays. This is true even though the objects are qualitatively identical (thus the effect is not feature based) and move locations after previewing (thus the effect is not location based). The effect is genuinely object based, hence the name OSPB. Kahneman et al. posited the existence of "object files" to explain the OSPB (as well as other aspects of object-based attention—see Kahneman and Treisman 1984). Object files are representations of particular objects that do not represent objects via any particular feature, but rather cluster features depending on whether they are instantiated in the same object.
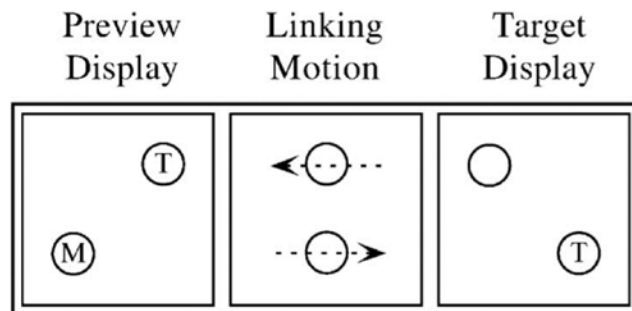
Figure 7—An object-reviewing paradigm (adapted from Mitroff et al. 2005)

Object files are linked to the "visual indexes" or "FINSTs" studied by Pylyshyn, Scholl, and others. While the notion of object files arose from studying the sometimes surprising ways features are clustered together in mid-level vision, the notion of visual indexes came from studying the

---

[26] Kahneman et al.'s (1992) original paradigm ask subjects just to name the letter rather than report a match, but since that version of the task doesn't actually require subjects to see the preview to perform accurately, the paradigm has since been modified to the one described in the main text.

equally surprising ways objects are tracked across space and time, unearthed in a fascinating series of studies regarding multiple-object tracking (MOT). In the classic MOT paradigm, participants foveate on a fixation cross while roughly eight objects (e.g., squares) populate the rest of the stimulus (Pylyshyn & Storm 1988). Some small number (e.g., four) of the objects will flash, indicating that they are to be tracked, while the rest are to be ignored. Participants are remarkably good at tracking three or four objects, even when the objects are qualitatively identical and intersect each other's paths and even when they are occluded by hidden barriers. Participants cannot even report the spatiotemporal criteria used to individuate and track objects—though, interestingly, if the motion ceases and an object disappears, subjects can say where it was and what direction it was headed in (Scholl et al. 1999), suggesting that "participants have conscious access to a spatio-temporal address of a currently attended object, though not always to other features of the indexed object" (Carey 2009, 75).

It is plausible that visual indexes and object files involve the same underlying capacity. Object files need spatiotemporal addresses of objects to maintain identity over time and across changes, which would be supplied by a visual index. Moreover, the OSPB is boosted under conditions of MOT (Haladjian & Pylyshyn 2008), which strongly substantiates the claim that visual indexes and object files are linked.

In order to avoid confusion, I will refer to perceptual object representations (PORs). PORs are representations of object that select individuals through the deployment of a visual index and store information about them (i.e., do the job of object files).

The mere existence of PORs places pressure on a wholly iconic model of perception. Recall that pure cases of iconic representation do not have a canonical decomposition into constituents. But PORs require a segmentation of the percept into constituent PORs, and thus rule out a naïve view of perception on which the visual system delivers a single pure, unsegmented icon.

This naïve view is obviously implausible, however, since other forms of segmentation like figure-ground segregation or other forms of perceptual grouping are incompatible with it as well. It would be uncharitable, therefore, to saddle anti-pluralists with this view. Instead, anti-pluralists can allow that the visual system segments the visual field, but still require that every output of perceptual segmentation processes is an icon. Carey (2009) seems to endorse a view like this about PORs (see Fig. 8).

Stimuli	Object-files (symbolic)	Object-files with properties bound (iconic)
	Object	
	Object Object	
	Object Object Object	
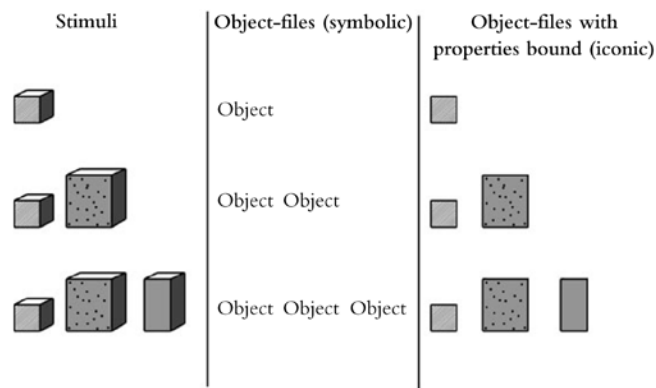
Figure 8—Carey's (2009) table distinguishing discursive and iconic models of PORs.

Carey points to a study (Feigenson, Carey, & Spelke 2002) in which infants were shown crackers of varying sizes deposited into different buckets and crawled toward the one with more total cracker, even if it was spread out over fewer crackers (e.g., one large cracker over two small crackers). The results "reveal the set-size signature of object-file representations" (Carey 2009, 146), which suggests that infants are using PORs with surface area bound to them (see also Block

unpublished). However, the data shows merely that infants sum over surface area as represented in PORs. Surface area could be represented and summed over whether encoded in an iconic or discursive format; it's not clear how the ability to sum over surface area bears on the format in which surface area is represented. Perhaps an iconic explanation is preferable all else equal because icons can more easily encode continuous variables like surface area. This argument is quite tentative, however, as Carey readily concedes (2009, 147).

There is a wealth of data that does shed light directly on the format of PORs. It strongly suggests that PORs are not icons.

4.2—Visual indexes. PORs incorporate an index-like representation of individual objects. The evidence for this hypothesis is that MOT works despite dramatic changes in features. For example, Zhou et al. (2010) found that changing color, luminous flux, and various shape properties did not disrupt MOT. Bahrami (2003) introduced a change detection task, where objects in an MOT paradigm moved around and became occluded by a "mud splash"—after the splash, either no objects changed or a single object changed its shape or color. Subjects often failed to register changes despite successfully tracking the objects, though detection of changes were better for tracked objects than untracked objects (see also Scholl et al. 1999).

All this evidence demonstrates that, while features are bound to objects in MOT, the binding is not holistic. The representation of the object can persist even though features are changing or lost entirely. For it to be possible to lose a feature while maintaining a representation of the object, the representation of the feature must be syntactically separate from the

representation of the object. This precludes holistic binding and instead requires a distinct constituent that stands for individual objects apart from various features. Thus PORs are not icons.

None of the foregoing argument requires that indexes are genuine mental analogues of bare demonstratives. Pylyshyn (2007) does hold that view, and it is indeed incompatible with an iconic model of PORs. But what matters is only that the binding between indexes and feature representations is not holistic. The data call for a representation of an object that comes apart from representations of color, luminosity, and various shape properties. This alone is sufficient to reject the view that a POR is an icon. An icon that encodes the color, shape, and location of an object does so holistically, and PORs lack this sort of holistic binding. This is true even if the indexical components of PORs always encode certain properties (e.g., location, trajectory, and even topological shape [Zhou et al. 2010]).

4.3—Abstract features. The fact that the indexical components of PORs are not iconic may be compatible with features' being encoded via an icon. The evidence tells against this view as well. Before looking at the evidence, however, note an *a priori* problem. This hybrid view holds that the indexical components of PORs are not iconic, and that featural components are; thus PORs comprise multiple formats at once. But how? If these representations are in genuinely distinct formats, how do they compose with one another? As mentioned above, different formats are akin to different languages, and should not ordinarily compose without some third language interpreting their relation to one another. The hybrid view therefore faces a serious *a priori* problem of unintelligibility.

This *a priori* point is not necessary to defeat the hybrid view, however. The object-reviewing paradigm can be used to determine how features are stored in PORs, which in turn can shed light on representational format. A series of experiments have showed abstract properties to be explicitly represented in PORs. Henderson (1994) found an OSPB for letters previewed in objects even when the letters differed in case from preview to test displays. Gordon & Irwin (1996) used entire words, e.g., the word 'bread' in all lower case appeared in the preview display and 'BREAD' in capitals was used in the test display. Given that there are no letterforms in common from 'bread' to 'BREAD', the identity of the word must be stored in a format that abstracts away from shape properties.

If features were represented iconically in PORs, then they should be bound together holistically. In that case, a POR should represent high-level properties like the semantic identities of words in a way that is bound up with representations of low-level properties—one should not find a high-level property represented independently of low-level properties. Tyler Burge, for example, who as aforementioned holds that perception is iconic, writes regarding face perception that "higher-level facial attributives never float free from low-level geometrical attributions" (2014a, 578). These object-reviewing results nonetheless show non-holistic encoding of high-level properties that float free from low-level properties.

Burge (2014a) appeals to a distinction between specific and generic low-level properties that ground perceptual attribution of high-level properties. For example, there might be a generic banana-shape property that the visual system uses to ground perceptual attribution of the kind

banana. One might therefore object that the studies just cited only show high-level properties floating free from specific low-level features, not generic low-level features.

A later study by Gordon & Irwin (2000) found the OSPB even when preview stimuli were words and test stimuli were corresponding pictures. For example, the previewed features might be the words 'apple' and 'bread' and the test stimulus would be a match if it were a picture of either an apple or a loaf of bread (see Fig. 9). Gordon & Irwin found that response times were quicker if the picture of the apple occurred in the same object in which 'apple' was previewed. There are no low-level features in common between the word 'apple' and a picture of an apple—not even highly generic low-level features. The POR needs to store a representation of the kind *apple* in a way that completely abstracts away from generic as well as specific low-level features. This capacity would be explicable if what's being stored is a distinct symbol that has the content <apple> and is bound to other features of the object in a separable, non-holistic fashion. That is, the capacity is explicable if the feature is encoded in a discursive rather than iconic format.
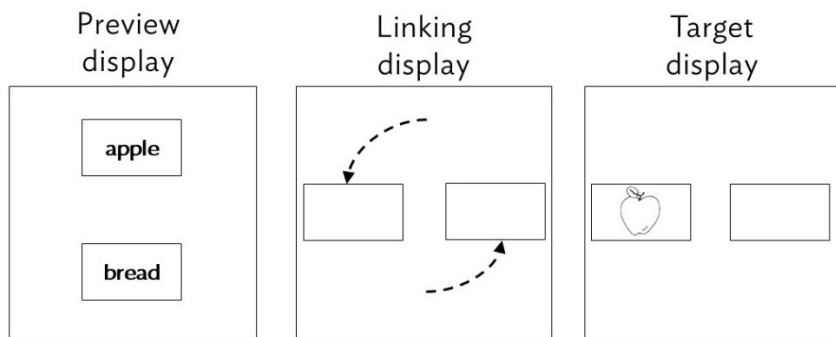
Figure 9—Illustration of Gordon & Irwin 2000

If discursive representations of categories are bound into PORs non-holistically, then we should expect kind representations bound into PORs in one modality to be accessible by another modality. This is precisely what Jordan et al. (2010) found. Jordan et al. used pictures as preview

109

features, after which objects would move to either side of the screen. Subjects then heard a sound come from one side and had to indicate whether the sound matched either of the previewed pictures. For example, suppose the previewed pictures were a telephone (in object 1) and a cat (in object 2), and 1 and 2 then move to the left and right, respectively. Hearing either a telephone ring or a meow from either side would be a match, but hearing a dog bark from either side would not. Jordan et al. found a OSPB in that subjects were quicker to correctly identify a match if it was heard coming from the same side as the object that had previewed the corresponding feature. To continue the example trial just mentioned, subjects would be faster if they heard a meow coming from the right than if it came from the left, since object 2 was previewed with a picture of a cat and moved to the right. The effect is not due to rehearsing lexical representations in verbal working memory, since subjects engaged in articulatory suppression (they repeated the words in their head during the study), thus loading verbal working memory.

This result vindicates the idea that discursive symbols that stand for categories and abstract away from low-level modality-specific bundles of features are stored in PORs. The authors draw this very conclusion, hypothesizing that PORs "store object-related information in an amodal format that can be flexibly accessed across senses" (2010, 500).

These results show that the thesis that PORs encode features in an iconic format is empirically untenable.

4.4—Low-level features. As a last retreat, one might admit that indexes and abstract categories are represented discursively in PORs but hold on to the idea that low-level features are represented iconically. This view preserves anti-pluralism in name only. It concedes that amodal discursive

110

symbols are deployed by the visual system alongside icons and therefore seems to be a version of perceptual pluralism. The desperate anti-pluralist could still insist that this view preserves the claim that every output of the visual system is at least *partly* iconic.

Unfortunately, even this view is not consistent with the evidence. If low-level features are represented in an icon in each POR, then low-level features are bound holistically and thus should not come apart from one another. As mentioned above, Bahrami (2003) found that subjects often failed to store information about the shape and color of tracked objects in an MOT paradigm (though not as often as they failed to store features of untracked objects). Bahrami also found differences in when shape and color were lost; for example, there was no difference observed in whether shape was stored in tracked objects between conditions with a mud splash occlusion and conditions without one, whereas there was a significant difference for color. Thus there were trials were shape was preserved and color lost in a POR.

Green and Quilty-Dunn (unpublished) argue that low-level features are not encoded iconically in PORs on the basis of independent feature loss in object representations in visual working memory. Their key example is an experiment by Fougnie and Alvarez (2011). Fougnie and Alvarez showed subjects an array of colored triangles at various orientations, then took them away long enough to ensure that they were stored in visual working memory; subjects then had to use a color wheel to change a probe item to the same color as the triangle in that location, and use an orientation wheel to make a probe item point in the same direction as the triangle in that location (see Fig. 10—the order of color and orientation responses was varied). Fougnie and Alvarez found that there were many trials on which color was stored and orientation was forgotten,

and vice versa. Green and Quilty-Dunn argue that PORs are stored in visual working memory and therefore that this result shows non-holistic feature binding, and thus non-iconic format, in low-level feature representations bound into PORs.
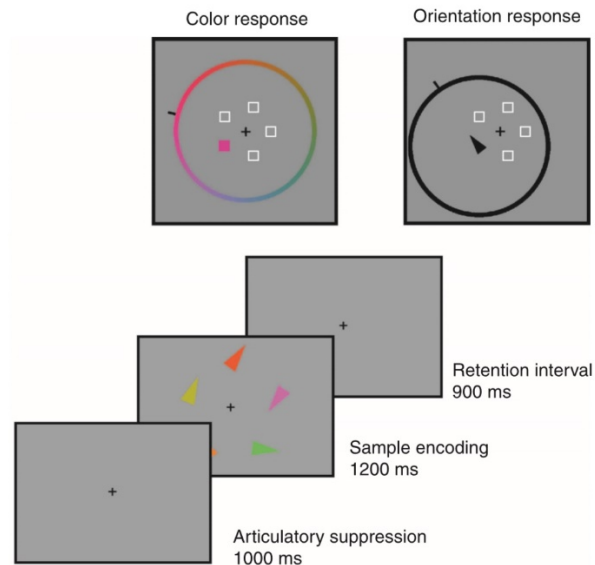
Figure 10—Independent feature failure in object representations in visual working memory (from Fougnie & Alvarez 2011)

There is no compelling positive evidence for iconic format in PORs, and a plethora of compelling evidence against iconic format in PORs. Combined with the arguments in favor of perceptual icons in the previous section, there is good reason to think that the visual system outputs both iconic and discursive representations. In other words, perceptual pluralism is true.

§5. Objections and Replies

I now turn to possible objections to the arguments in the preceding section.

5.1—PORs are not really perceptual. Perhaps the most tempting immediate response is to reject the basic assumption of my arguments concerning PORs, viz., that PORs are outputs of the visual

system. In that case, the moniker 'POR' is a misnomer; thus for the rest of this subsection I will instead use the neutral term 'object file'.[27]

Block (unpublished) pursues this strategy. According to Block, the outputs of perception are iconic while object files are post-perceptual representations that are proprietary to working memory. Block does not explicitly concede that the arguments from Quilty-Dunn (2016) and Green and Quilty-Dunn (unpublished) establish that object files are discursive, but his main line of response is to insist that object files are post-perceptual rather than that they are iconic.

There is no positive evidence that object files are post-perceptual, however, and there is substantial reason to think that they are genuinely perceptual. Since Quilty-Dunn (2016) and Green and Quilty-Dunn (unpublished) simply assume that object files are perceptual without extensive argument, I will address this question here.

The view that object representation is always post-perceptual was famously defended by Elizabeth Spelke:

> Human perceptual systems appear to analyze arrays of physical energy so as to bring knowledge of a continuous layout of surfaces in a state of continuous change. We perceive the layout and its motions, deformations, and ruptures. This continuous layout contains no spatially bounded "things" and no temporally bounded "events": Perceptual systems do not package the world into units.
>
> (Spelke 1988, 229)

Spelke's argument, however, simply assumes that perceptual representations are wholly iconic. Her argument would rule out *any* form of segmentation as being genuinely perceptual, and therefore seems to assume an overly impoverished view of perceptual representation. Moreover,

---

[27] As mentioned above, I use the term 'POR' to link object files and visual indexes together, whereas the term 'object file' does not do that; the connection between object files and visual indexes will presently become significant.

much of what drives Spelke's argument is the fact that object files are accessible across modalities and represent relatively high-level properties like solidity. Similarly, Carey argues that object files are in "core cognition" rather than perception (though see Carey 2011) because they "cannot be stated in the vocabulary of perception" (2009, 63) since they do not reduce to "perceptual or sensori-motor primitives" (2009, 67). But the thesis that object files have a discursive format can explain these properties without positing that these representations are in any sense post-perceptual.

Carey also argues that object files have a "rich conceptual role" (2009, 94). She cites evidence that "infants as young as 2 months old represent physical relations between objects such as inside and behind, and their representations are constrained by knowledge of solidity—a property of real objects but not of 2-D visual objects" (2009, 103). She also mentions that infants expect objects to be "subject to the laws of contact causality" and that they "represent objects as the goals of human action" and "represent self-moving agents as the cause of motion of inanimate objects" (2009, 103). She concludes that object files

> are inferentially interrelated with other representations that themselves cannot be reduced to sensory primitives, other representations that are the outputs of different core systems and of domain-general learning processes. Thus, they play a central conceptual role, one of the hallmarks of nonmodular, conceptual processes.
> (Carey 2009, 103)

Some of this data seems to be explained in terms of the fact that object files include representation of properties like solidity that influence looking times and other behavioral measures. The fact that an object file can encode these properties and influence behavior in doing so does not in itself require a central conceptual role (though it arguably would if one assumed that

all perceptual representations must reduce to transduced primitives). Other data does seem suggestive of genuine inferential transitions—I will return to that issue in Chapter 3. For now, however, note that in all these cases object files function as *premises* in inferences, not *conclusions* of inferences. A representation might wholly belong to the visual system and yet be fed into cognition to function as a premise in inferences.

Any representation with propositional structure will be apt to function as a premise in inferences. An argument that a representation that functions as a premise in inferences cannot also be perceptual would therefore rule out the possibility of propositional structures in the visual system by fiat. The question of whether there are propositional structures in the visual system does not seem decidable in this way, however; it seems like a live empirical question (see Chapter 3). The fact that object files may feed into cognition in this way thus does not undermine the possibility that they are genuinely and fully perceptual representations. Moreover, recall the discussion of the translation problem in Chapter 1. It is a benefit of architectural approaches to the perception–cognition border that they deliver representations in the same format as cognition, which can then be used directly by cognition. On this conception of the relation between perception and cognition, at least some perceptual representations *ought* to figure as premises in inferences. I conclude therefore that the highly constrained "conceptual role" of object files does not suffice to show that they are post-perceptual in any sense.

Block instead relies on the fact that object files are stored in visual working memory (VWM) to undermine the claim that they are constructions of the visual system. It is not obvious from the basic object-reviewing paradigm that initially established the existence of object files that

they are stored in VWM. The operationalization of representations stored in VWM involves, first, the disappearance of the object during a retention interval long enough to ensure that the representation makes it into VWM. Some kind of task is then performed (e.g., a change detection task, wherein an object may or may not change a feature like color or shape and the task is to indicate whether there was such a change). Successful performance on this task after a sufficient retention interval ensures that the representation was stored in VWM.

In object-reviewing experiments, on the other hand, objects are typically in plain sight for the duration of the trial. This aspect of the experimental design reflects the fact that object files were not originally posited as items to be stored in VWM. Instead, they were posited to explain proprietarily visual phenomena, and were considered "mid-level" constructs of the visual system, after the most basic feature detection but prior to late vision or post-perceptual processing (Kahneman et al. 1992). Indeed the seminal study of object-based representations in VWM was not done until years later by Luck and Vogel (1997).

A well-known problem in vision science concerns the correspondence of two temporally contiguous and qualitatively distinct retinal inputs (Ullman 1979). Object files were posited partially to solve this correspondence problem by allowing changes in retinal input due to (i) the movement of objects, (ii) changes in the features of objects, or (iii) saccades, i.e., movement of the eyes, to be coherently integrated by appeal to representations of enduring objects (Kahneman et al. 1992, 179). Segmenting the world into coherent, enduring units that can gain, lose, or change features while retaining their identity allows visual processing to make sense of retinal input.

Without object files, vision would be as William James imagined it to be for infants, a "blooming, buzzing confusion."

Moreover, object files were posited also to be the vehicles of object-based attention. In Treisman's feature integration theory, features are bound via object-based attention, which led to the positing of object files by Kahneman and Treisman (1984) and then the discovery of the OSPB by Kahneman et al. (1992). The OSPB is itself an instance of object-based attention.

Object correspondence, feature integration, and object-based attention are all properly within the scope of vision science. These are phenomena of online visual processing rather than post-perceptual processing and manipulation in working memory. The core theoretical roles of object files thus tell against a post-perceptual view.

The main point just made is that storage in VWM is not essential to object files and was not why they were posited in the first place. There is nonetheless some evidence that object files are entered into VWM. Hollingworth and Rasmussen (2010) simply combined the basic object-reviewing paradigm with the basic VWM change detection paradigm. Four objects were presented, and colors were previewed in them; then the features disappeared, 300ms passed, and the objects visibly moved to other locations for 500ms and were still for 200ms; next, one second now having passed since the preview phase, colors were shown in the objects again (see Fig. 11). In the "updated" condition, the colors followed the objects; this is the important condition for testing the role of object files in VWM. In the "original" condition, they were shown in their original locations (i.e., in different objects at the same location). In the "no correspondence" condition, they were shown in different objects at different locations. Either all colors were the same or one color had

changed. The change detection task—a standard way of probing the contents of VWM—required subjects to say whether or not any colors had changed.
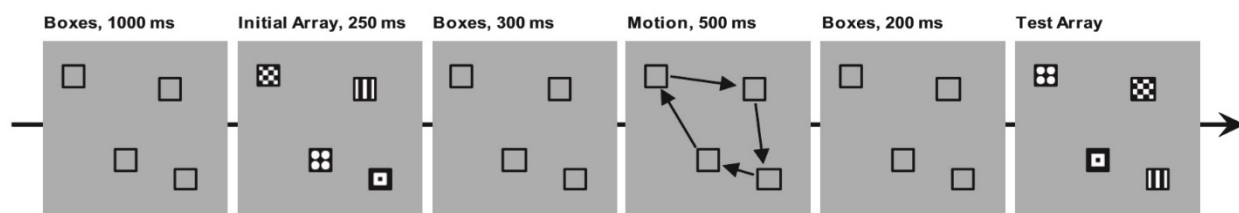
Figure 11—From Hollingworth & Rasmussen 2010

Hollingworth and Rasmussen found that accuracy and reaction time on the change detection task were better for the updated condition than the no correspondence condition. That is, performance on the change detection task was facilitated by the use of object files. Since change detection after a 1000ms interstimulus interval is a standard way of probing VWM, there is good reason to think that object files can be stored in VWM.

There is, however, no reason to conclude from this that object files are not constructed by proprietarily visual processes. Manipulation of representations in VWM is an instance of cognitive processing. But it does not follow that *storage* of representations in VWM is an instance of cognitive processing. And it certainly does not follow that representations stored in VWM are *formed* through post-perceptual cognitive processes. On the contrary, it is much more plausible that representations that are formed by the visual system can be entered into VWM, and once in VWM can be cognitively manipulated. Thus Block's inference from the fact that object files can be stored in VWM to the claim that object files are not formed by the visual system is a *non sequitur*.

The hypothesis that object files are only formed after visual processing has finished raises some odd questions. Are there any representations of objects delivered by visual processing? If so, how do we study them experimentally? Expelling object files from vision science—and with them,

object-reviewing paradigms—robs us of basic experimental tools for studying object perception. What explains visual correspondence? What are the vehicles of object-based attention? If object files are perceptual, these questions have satisfying answers that are unified by the same theoretical construct. If object files are post-perceptual, then there is no clear answer to any of these questions.

There is a related dilemma for the view that object files are post-perceptual. Either there are perceptual object representations or not. If there are not, then none of the questions in the previous paragraph are answered. Furthermore, it is extremely implausible both phenomenologically and empirically that there are no perceptual representations of objects. MOT, for example, seems clearly to qualify as a visual phenomenon both intuitively and scientifically, as do object-based attention and the solution (whatever it is) to the correspondence problem.

If there are perceptual object representations, then what are they like? How do we know about them? One might appeal to MOT. But MOT boosts the OSPB in object-reviewing paradigms, strongly suggesting that the vehicles of MOT (i.e., visual indexes) are constituents of object files (Haladjian & Pylyshyn 2008). This horn of the dilemma thus requires either rejecting that MOT is genuinely perceptual, or providing an alternate explanation of why MOT and object-reviewing seem to be so intimately related. Rejecting that MOT is genuinely perceptual would seem to eliminate any way of studying perceptual object representation. And an alternate explanation of why MOT boosts object-reviewing may be forthcoming, but has not yet been offered. Moreover, storage of features of objects happens in MOT (Bahrami 2003), and the relation between MOT and the OSPB provides reason to think that this is via object files.

The view that object files are post-perceptual is thus unmotivated and faces serious theoretical problems. But there is also more positive evidence that object files display hallmarks of perceptual processing. These hallmarks are (a) being informationally encapsulated from cognition, and (b) being integrated into clearly perceptual processes.

One form of evidence for the informational encapsulation of object files is the divergence between cognitive and perceptual criteria for object individuation and tracking. For example, most people know that an object that shrinks down before it goes behind an occluder does not cease to be the same object, and should tend to continue along the same trajectory. Nonetheless, when an object shrinks before heading behind an occluder, tracking is severely disrupted (Scholl & Pylyshyn 1999). For another example, most people know that adding a line (say, a piece of string) from one object to another does not make them the same object. Nonetheless, when a line connects two objects, tracking is severely disrupted (Scholl, Pylyshyn, & Feldman 2001). These are clear cases of informational encapsulation; the object perception system operates on its own proprietary store of information that excludes information stored in central cognition. Moreover, one can view the relevant demos for oneself and consciously experience the effect.[28] This encapsulation also goes in both directions, since the information used for visual tracking is not accessible to cognition. It is a surprising result that visual tracking is limited in these ways, not a mere verification of something we already (cognitively) knew to be the case.

Another, even more striking case of informational encapsulation was found by Mitroff et al. (2005). Mitroff et al. ran a typical object-reviewing paradigm, except the motion of the objects

---

[28] Demos are available on Brian Scholl's website: http://perception.yale.edu/Brian/bjs-demos.html

intersected in a way that was visually ambiguous between two objects "bouncing" off one another and two objects "streaming" through one another (see Fig. 12). They ran an object-reviewing study where features were previewed before the ambiguous motion and test features were displayed afterward. They also asked subjects to judge whether the objects had bounced or streamed. Remarkably, they found a sharp divergence—the OSPB showed that the object files had bounced even though subjects judged the objects to have streamed through each other. This result is another clear example of informational encapsulation. The processes that output object files are stimulus driven and informationally encapsulated, and thus seem to be genuinely perceptual.
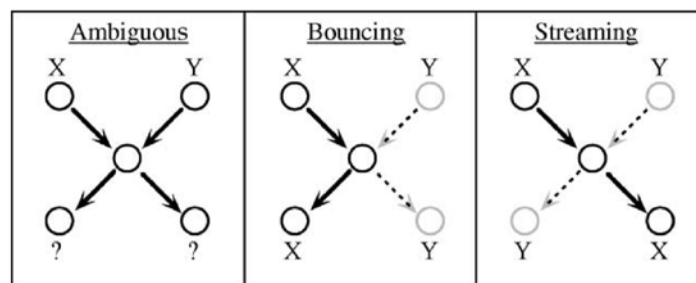
Figure 12—From Mitroff et al. 2005

Another way of arguing that a representation is genuinely perceptual is to show that it is integrated into perceptual processes. One example of this in object files is the demonstrated relationship between object files and visual indexes (Haladjian & Pylyshyn 2008). Moreover, object-based attention warps the perceived spatial properties of objects (e.g., Vickery & Chun 2010; Liverence & Scholl 2011). Given that object files are plausible candidates for the vehicles of object-based attention, object files are therefore integrated into spatial perception.

Another perceptual process which operates on object files is the guiding of saccades. Upon walking into an unfamiliar room, your eyes might wander around and thus "saccade" to various parts of the room. Perceived information about the scene is used to guide where we saccade and to

maintain a stable percept across saccades. This phenomenon is undeniably perceptual. It is hard to imagine what it would be like to visually perceive if the visual system did not maintain information across saccades—our percepts would be an incoherent series of snapshots.

Episodic information used to guide saccades is referred to as "transsaccadic memory" (Irwin & Gordon 1998). Transsaccadic memory is exploited in low-level visual tasks like tracking a moving object. For example, if an animal is in your periphery and you move your eyes to focus on it, the animal might have moved in the time it took you to saccade; it is nonetheless quite easy to perform a second, "corrective saccade" and come to fixate on the animal at its new location. While the guidance of saccades is often (though not always or even usually) under voluntary top-down control, the *information stored about the scene* is perceptual and is used in perceptual processes.

There is evidence that object files are stored in transsaccadic memory (Irwin & Andrews 1996; Irwin & Gordon 1998). For one thing, transsaccadic memory does not seem to tap into iconic memory. Irwin (1992) presented subjects with an array of letters at one fixation and, after saccading, presented a partial report cue to a subset of letters, similar to the Sperling (1960) experiments discussed above. Unlike in the Sperling experiments, subjects only showed storage of three or four letters. This suggests that iconic memory is not used for transsaccadic memory (and indeed suggests that icons are not stored in transsaccadic memory).

Schut et al. (2017) argue that "corrective saccades are executed on the basis of object files" (2017, 138). They presented subjects with six objects around a fixation point and instructed them to saccade to whichever object briefly expanded in size and then return fixation to the center (see

Fig. 13). A second target object (which may or may not have been the same object) would then briefly expand, cuing subjects to saccade to it. In some trials, the objects would rotate during the saccade so that fixation landed next to the second target instead of on it, in which case subjects had to initiate a corrective saccade. Schut et al. found that reaction time was quicker in conditions where the second target was the same object that had previously been fixated (i.e., the first and second target are the same object). They found the same effect when objects changed color after fixation and when objects changed location after fixation (by slowly rotating between the first saccade and the cuing of the second target object—not pictured in Fig. 13). The effect is therefore genuinely object based and taps into the same object representations that are used to track correspondence of objects across changes in features and locations, viz., object files.
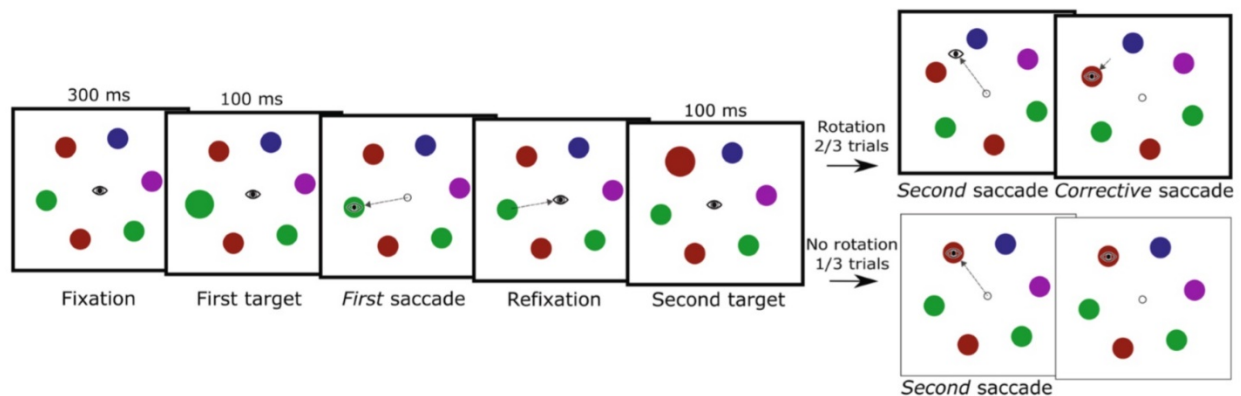
Figure 13—From Schut et al. 2017

Like object-reviewing paradigms, tests of transsaccadic memory can shed light on representational format. Pollatsek et al. (1984) used a basic paradigm where subjects would fixate on the center of the screen with an object in the periphery, then saccade to the object and name it as quickly as possible. Changing the features of the object during the saccade slows down the ability to name it, showing that features are encoded in transsaccadic memory. Pollatsek et al. switched

out objects with a different exemplar of the same basic-level category, with sometimes substantially different low-level features (e.g., a young cat with textured fur facing forward vs. an older cat without textured fur facing to the left, or dogs of two different breeds—see Fig. 13). They still found significant facilitation in naming the object despite the change in low-level features.

Figure 14—Stimuli used by Pollatsek et al. (1984)

Rayner et al. (1980) used a similar paradigm with words instead of pictures, and found that words could vary in case across saccades and still facilitate naming. Pollatsek et al. (1990) used the same stimuli as Pollatsek et al. (1984) (e.g., the two visually dissimilar cats), but varied the location of objects pre- and post-saccade. They still found a significant facilitation for naming objects across saccades despite the change in location and low-level features. This suggests that the conceptual identity of the picture is encoded in transsaccadic memory without holistic binding to low-level features.

There is therefore evidence for representations in transsaccadic memory encoding kinds just as there is for object files. This bolsters both the claim that object files are stored in transsaccadic memory and the claim that perceptual processes utilize discursive kind representations.

Objection: These transsaccadic memory effects only show storage of what Burge calls "generic" low-level features, like a generic cat shape.

The objection is correct in that (e.g.) Pollatsek et al. (1984) did not control for this alternative. But the independent evidence that object files are stored in transsaccadic memory, together with the independent evidence that object files encode kinds independently of generic low-level features, should bias our interpretation of these results in favor of the claim that representations in transsaccadic memory encode kinds. Moreover, there is evidence that which low-level features are used to guide object representations in transsaccadic memory is determined by which kinds are represented rather than merely the presence of the feature.

Gordon and Vollmer (2010) had subjects saccade to a location between two objects, at which point the two objects were replaced by a single object. The task was to name this object. The object either did or did not match the category of one of the two previewed objects, and may or may not have changed color. Gordon and Vollmer found an object-specific effect: a decrease in reaction time for naming the object above(/below) the fixation point if that object was previewed above(/below) fixation. This effect was significantly reduced when color changed, *but only for objects with diagnostic colors*. Thus the correspondence between transsaccadic representations of a banana depended on the yellow color of the banana, whereas correspondence for transsaccadic representations of an object without a diagnostic color (such as a pail) did not depend on its color. This result suggests that the representation is not maintained merely on the basis of low-level properties; which low-level properties are used for object correspondence depends on their relation to conceptual kinds that are explicitly represented in transsaccadic memory (see also Gordon et al. 2008; Gordon 2014).

The fact that object files are used in transsaccadic memory bolsters the key hypothesis that led Kahneman and Treisman (1984) to posit them in the first place: object files are used to render otherwise shifting visual stimulations into coherent percepts of scenes of objects that maintain their integrity across changes in their features and where they are located. Pushing object files outside the purview of visual processing renders the coherence of visual phenomenology mysterious.

We should conclude on the basis of this discussion that object files are genuinely perceptual object representations; thus I will continue to use 'POR' in what follows.

5.2—The Berkeleyan response. One might grant that object representations are perceptual but seek to explain results that push toward pluralism by giving icon-based explanations. One of the key arguments from Section 4, as well as the discussion of transsaccadic memory in Section 5.1, concerns representation of abstract kind properties. Abstract contents have also long been a bugbear of empiricist theories of concepts—arguably for the same reason, viz., the proposed iconic format of the representations. A Berkeleyan response is to argue that icons can be *used* in ways that are indicative of abstract kinds without explicitly representing those kinds via a distinct symbol that abstracts away from low-level, iconically representable properties.

Block (unpublished) argues that a Berkeleyan response is sufficient to ward off the argument from abstract kind contents to discursive format in perception; to suppose that icons have to represent maximally determinate properties is to commit the "photographic fallacy" (Block 1983). There may be ingrained functional roles for perceptual icons, Block argues, that allow an

icon to fulfill a role that a pluralist would explain in terms of discursive format. He suggests that some of these roles are biologically preferred.

This Berkeleyan response is not plausible. Berkeley's preferred example was the use of mental image of a triangle (which must be either equilateral, isosceles, or scalene) in a way that functions as a representation of triangularity in general. One Berkeleyan idea is that one can attend to some features of the mental image and ignore others in a way that privileges uses that pertain to the attended features, thus simulating abstraction from the ignored features without eliminating them from the representation (Prinz 2002, 29).

But the data on abstract features in object files cannot be explained this way. Consider again Gordon and Irwin's (2000) finding of an OSPB despite the preview being the word 'apple' and the test feature being a picture of an apple. The problem is not that an iconic representation of an apple functions in respect of some iconically represented features and not others. The problem is that the representation of an apple in an object file is not holistically tied to any low-level features at all. There is nothing iconically representable in common between the word 'apple' and a picture of an apple.

The argument for pluralism here is not that it is logically impossible that the information in the POR is encoded iconically. Rather, the argument is that an iconic model of PORs offers no explanation of this effect, whereas a discursive model nicely explains it. There's no reason why an iconic representation of an apple should play the role that the POR does in Gordon and Irwin's (2000) study. Replying that the icon *just does* play that role restates the data rather than explaining

127

it. The claim that the role is biologically ingrained is also not plausible for kinds like *apple*, *pencil*, or *piano*, which are encoded in PORs.

Block (unpublished) argues that discursive representations also need specific functional roles in order to have content, including to represent kinds, so they are no more explanatorily successful than icons. But this reply is unsuccessful. Discursive representations always have some particular interpretation in the systems in which they are deployed. The fact that discursive representations need not syntactically correspond to what they represent or holistically encode features in the way icons do frees them up to explicitly represent a wider variety of properties, including abstract ones (i.e., properties that are not holistically tied up with particular low-level features, or that don't have spatiotemporal parts). The problem for empiricism in general and for iconic models of PORs in particular is that icons on their own typically do not explicitly represent anything more than low-level features. An icon is typically compatible with multiple high-level interpretations. Cases where the high-level interpretation is settled in one way rather than the other are best explained by positing discursive representations that specify the relevant interpretation. The Berkeleyan response, therefore, does no more than restate the data and offer a logical possibility that icons *just do* play the right role whereas a pluralist view provides a substantive explanation.

5.3—That's not what icons are. One might simply reject the characterization of icons used in this chapter. Perhaps some other notion of iconicity can avoid these objections and explain the data on object perception.

This objection is idle unless a notion of iconicity is offered and given independent substantiation. Section 3 of this chapter contained a great deal of empirical substantiation of the notion of iconicity used here, viz., a correspondence between parts of the representation and parts of the scene or object, and the holistic encoding of features. Moreover, the notion of iconicity used here has been used throughout cognitive science and is at the heart of Kosslyn's (1980; 1994) theory of mental imagery, which even an opponent like Pylyshyn says "can be taken as the received view in much of the field" (2002, 181).

Block (unpublished) writes that nobody has adequately defined iconicity. One could quibble with this claim, given the detail and explanatory success of Kosslyn's model of mental imagery. But even if nobody has adequately *defined* iconicity, the *phenomena* that are explained by iconicity such as mental imagery, iconic memory, ensemble perception, and low-level spatial perception indirectly provide a characterization. I argued above that a notion of iconicity that appeals to parts that represent parts and holistic binding of features and individuals explains all this data and therefore should form the basis of any theory of iconicity (which I'll try to provide in Chapter 4, ICONIC REPRESENTATION). PORs do not have these properties. PORs therefore are not iconic in the explanatorily useful sense of that term.

The opponent who rejects this notion of iconicity takes on a dialectical burden of articulating another notion that can explain mental imagery, iconic memory, and other key phenomena that perceptual icons are supposed to explain and showing how that notion of iconicity can also explain perceptual object representation. Since this burden has not yet been met, this objection is merely promissory.

§6. Perceptual Pluralism as a Research Program

The visual system delivers perceptual icons as well as discursive PORs. Therefore, perceptual pluralism is true. There are plausibly other representations outputted by perceptual systems that are neither perceptual icons nor PORs. I will briefly discuss possible examples. The evidence in these cases is less decisive than the evidence for PORs, but it is probative nonetheless.

6.1—Face perception. Any application of a high-level category to a distal stimulus will likely involve segmentation of early iconic perceptual representations and thus may involve the tokening of a distinct discursive representation that functions to indicate the relevant category. One possible example is face perception.

Whether faces are literally perceived is a matter of some controversy, but faces are one of the best empirically grounded candidates for high-level perceptual content to date (see Bruce and Young 2000 and Little et al. 2011 for overviews), with a possible dedicated neural apparatus in the fusiform gyrus subserving a module for face perception (Kanwisher et al. 1997). Block (2014) has recently argued that faces are genuinely perceived on the basis of perceptual adaptation effects for faces (Butler et al. 2008, Susilo et al. 2010). I'll assume for the sake of argument that there are "perceptual attributives for facial expressions" (Block 2014, 563) and other aspects of face perception.

The visual system must have some mechanism(s) that take iconic perceptual representations (perhaps with some non-facial segmentation, e.g., into visual objects) as inputs and delivers representations of faces as outputs. Face perception is a multi-stage process. Prosopagnosia is a condition marked by the breakdown of a specific type of recognition pertaining to faces.

Prosopagnosics show qualitatively and quantitatively different levels of impairment at different stages of processing, with some unable to differentiate faces at all, and others able to differentiate faces but unable to access memories for purposes of recognition (Sergent and Signoret 1992). Prosopagnosia is typically understood as a failure of facial *recognition*, the capacity to distinguish between and reidentify faces, but a "preliminary and prerequisite" ability is face *detection*, i.e., the capacity to discriminate facial from non-facial stimuli (Lewis and Edmonds 2003, 903). Faces are identified on the basis of low-level features such as holistic triangular configurations of eyes and mouths (Lin and Fan 2001).

The fact that face perception is triggered by holistic configurations suggests that icons are *inputs* to face processing, but perhaps at least some of the *outputs* of face processing (i.e., perceptual representations of faces as such) are discursive. It may be that face detection involves tokening a discursive representation with the content <face> that is not holistically bound up with low-level features. Face recognition might operate on discursive face detection and iconic representations of low-level features to deliver as output a discursive tag that marks the face as belonging to a particular person (which can then serve as inputs to mechanisms that store the particular face representation in memory and search memory for instances of the same face, etc.). Since prosopagnosia is not associated with low-level perceptual deficits, prosopagnosics may have the same early iconic representations as normal perceivers but lack the capacity to construct discursive face representations (i.e., failures of detection) or to store or check those representations against memory (i.e., failures of recognition).

It is certainly not impossible that face perception outputs icons rather than discursive representations. However, a discursive view predicts the dissociation of prosopagnosia from low-level perceptual deficits, since it posits a discrete symbol standing for the property of being a face (as well as the identity of an individual face) that is not holistically bound to low-level features. Failure of the capacity to represent faces need not involve failure of the capacity to represent low-level features if the two capacities involve distinct representational vehicles. An iconic view can posit distinct mechanisms for faces and for low-level features to explain this dissociation, but does not seem to readily predict the distinction the way a pluralist view does.

6.2—Ensemble perception. I earlier argued that ensemble perception provides evidence for perceptual icons. Now, however, I want to argue that it may also provide evidence for discursive representations in perception.

The reason to think ensemble perception suggests the existence of perceptual icons is that the *inputs* to ensemble perception are iconic since they do not segment out each individual item in the ensemble. But what about the *outputs* of ensemble perception, i.e., the ensemble representation itself?

Prima facie, ensemble representations seem to violate the holisticity constraint on icons. Ensemble perception abstracts away from other features and spits out a representation of an average along a particular property dimension. All else equal, one would expect a representation that operates on multiple properties of icons to respect the holisticity of those icons and thus operate on those properties together. Ensemble perception, however, does not work this way.

Indeed Hubert-Wallander and Boynton (2015) found evidence that ensembles of different properties are computed in very different ways. They presented subjects with a series of circles one at a time of varying sizes and at various locations on the screen. Subjects were prompted before each trial to indicate the average size or the average location of the set. Hubert-Wallander and Boynton examined how subjects weighted each individual object in the series. They found that earlier objects were weighted more in the derivation of average position while later objects were weighted more in the derivation of average size. These results "imply that different mechanisms are associated with different feature domains" (Hubert-Wallander & Boynton 2015). While this is not logically incompatible with the claim that ensemble representations are iconic, the use of separate mechanisms to derive separate feature ensembles is predicted by a view on which ensemble representations are discursive and thus non-holistic.

Furthermore, there is evidence that ensemble representations, like PORs, are entered into VWM and exhibit the same set size limitations. Im and Chong (2014) showed subjects multiple sets (up to five sets) of circles of varying sizes at once. Sets were delineated by color, though in one condition sets were also spatially grouped together. Subjects then indicated which of two probed sets had a larger average. Im and Chong found that "participants correctly remembered at least 2.5 sets without [spatial] grouping and 3.5 sets with [spatial] grouping" (2014, 669). This limitation held whether or not there was a one-second retention interval before the probe, suggesting that the effect taps into VWM. Ensemble representations are entered into VWM and exhibit the same set size limitations as PORs. This effect is predicted by the pluralist view that ensemble representations

are discursive and thus display the same sort of "item effect" (Fodor 2007) as other discursive representations like PORs.

6.3—Going forward. There are potentially many other examples of non-iconic format in perception. For example, various forms of perceptual grouping and segmentation might turn out to be best explained in terms of non-iconic format(s). E.J. Green (forthcoming) argues that the visual system deploys structural descriptions of part-whole structures of objects that seem to require a canonical decomposition and some degree of abstraction away from low-level features. These structural descriptions might be genuinely perceptual and yet non-iconic.[29] Moreover, Potter's (2014) work on the apparently lightning-fast deployment of concepts during rapid serial visual presentation might push in favor of perceptual mechanisms that deploy discursive concepts (Mandelbaum forthcoming).

Rather than investigate these possibilities further here, I'll simply note that the truth of perceptual pluralism opens up these avenues for further research and makes more plausible the claim that these cases involve discursive or even *sui generis* representational formats.

§7. Conclusion

The description and motivation of perceptual pluralism in this chapter was all pitched at the level of representational format. In particular, I argued that perception delivers both iconic and discursive formats.

---

[29] Green's (2015) work on layered shape perception and the perceptual representation of affine shape in addition to metric shape (e.g., *parallelogram* in addition to *square*) also arguably makes more sense on a pluralist view than an anti-pluralist one.

However, perceptual pluralism proper is the thesis that perception delivers proprietary representations as well as tokens of the same type used elsewhere in the mind, especially including those used in thought. One might thus resist a fully pluralist interpretation of the evidence cited above by arguing that, while PORs and thoughts might both have a discursive format, the former are not genuinely *conceptual*. Thus object perception does not involve tokens of the same types of states deployed in thought. In the next chapter, therefore, I will argue that PORs are not only discursive but also conceptual and propositional. This further claim is necessary to defend perceptual pluralism in the strongest sense.

We do nonetheless have reason to reject the most prominent version of anti-pluralism, according to which the outputs of perception are exclusively iconic. Moreover, we cannot draw the border between perception and cognition by appeal to the hypothesis that percepts are iconic and thoughts are discursive. It is thus not plausible that the perception-cognition border is based on differences in representational format.

SEEING OBJECTS: A CASE STUDY IN THE UNDER-INTELLECTUALIZATION OF PERCEPTION

Chapter 3

"Let me see if I've got this straight, Prof. Skinner: you're saying it's a mistake to anthropomorphize <u>humans</u>?"

—*Sydney Morganbesser (as quoted by Daniel Dennett)[30]*

§1—Thinking and Perceiving

1.1—Representations and mental architecture. As discussed earlier, there are at least two broad sorts of ways of differentiating perception and cognition: first, *representational* strategies posit distinct types of representations in perception and cognition; second, *architectural* strategies posit psychologically primitive functional borders between perception and cognition. Representational strategies argue that the kinds of states deployed in perceptual systems are internally unlike those in cognitive systems. According to one recently prominent form of the representational strategy, perceptual representations are iconic while cognitive representations are discursive (e.g., Carey 2009; Burge 2010a, 2014; Block 2014).[31] I argued in the previous chapter that this sort of

---

[30] Dennet 2006, 109; emphasis his.

[31] I put aside questions about whether the *content* of perceptual states is distinct from the content of cognitive states (e.g., Peacocke 1992, 2001; Stalnaker 1998; Heck 2000; Byrne 2005; Speaks 2005; Burge 2010a). From the perspective of cognitive science, a difference in content will only make a difference if it is reflected in the state itself. It is thus *state* nonconceptualism rather than *content* nonconceptualism that matters for pursuing the representational strategy. It is

representational strategy fails due to the empirical fact that perceptual systems output discursive

representations such as perceptual object representations (PORs).

One might jettison the idea that perceptual representations are wholly iconic and insist

instead that PORs are discursive but still nonconceptual representations that are in some way

internally unlike the conceptual states deployed in thought. The conclusion of Chapter 2 is indeed

neutral on this question. Once we establish that PORs are discursive, there are two possibilities:

either they are conceptual representations (and the representational strategy fails) or they are

nonconceptual representations in a discursive format.

In the absence of an independent account of nonconceptual discursive representations in

perception, the natural hypothesis would be that discursive representations in perception are the

same as those that occur in cognition. Even if there is a reason to posit nonconceptual but

syntactically structured representational formats in perception (perhaps Biederman's [1987] geons

or Marr's [1982] viewpoint-invariant three-dimensional models; see also Green forthcoming),

furthermore, object perception in particular may turn out to display hallmarks of conceptuality.

My aim in this chapter is to take it as a working hypothesis that PORs are conceptual and

see whether that hypothesis suffers any serious objections. I will run through a multitude of

conditions for conceptuality, often modifying them as I go. My conclusion is that, once clarified, all

appropriate conditions for conceptuality are met by PORs. Perceptual systems deploy conceptual

representations in propositional structures to represent individual objects as such and attribute

useful to read this form of the debate as at least partly a debate about representational format (see Chapter 1 for
discussion).

features to them. If there is a border between perception and cognition, therefore, it must be architectural.

1.2—How to intellectualize perception. Philosophers commonly accuse one another of "over-intellectualizing" various mental phenomena. This accusation is nowhere more common than in the philosophy of perception. Robert Audi, for example, argues that "we cannot properly understand perception if we over-intellectualize it" (2015, 74; see also Nanay 2013, 9–10). Tyler Burge complains that "nearly all prominent philosophical work on this topic over the previous century over-intellectualized" conditions for objective perceptual representation (2010a, xi). Successful avoidance of over-intellectualization is sometimes cited as a reason to abandon representational approaches to perception altogether (e.g., Noe 2015, 1).

There are at least two ways in which one might intellectualize perception: one might intellectualize perceptual *processing*, and one might intellectualize perceptual *states*. Claiming that perceptual processing is fundamentally the same as reasoning would be a case of intellectualizing perceptual processing. One could argue, for example, that perceptual processing involves the same exact forms of inference exhibited in rational thought, such as *modus ponens*; or one could argue that perceptual processing has access to background beliefs; or one could argue that perceptual processing requires intellectually demanding representation of the conditions of objective representation (see Burge 2010a, chapters 5, 6, and 7, for historical discussion and criticism). Intellectualizing perceptual states, on the other hand, would be a matter of claiming that perceptual states are conceptual in just the same way that states involved in reasoning are conceptual. It might

be that perceptual states, just like thoughts, involve the deployment of concepts in propositionally structured, truth-evaluable mental representations.

It is important not to run these two forms of intellectualism together. In particular, intellectualism about perceptual states—what we might call *conceptualism*—does not require intellectualism about perceptual processing—what, in one form, we might call *inferentialism*. One can endorse conceptualism, and hold that perceptual states are conceptual, without supposing that the perceptual processes that generate them operate according to rules of inference, have access to background knowledge, or depend on knowledge of conditions for objective representation.

There are compelling reasons behind the now-orthodox rejection of unreconstructed inferentialism (a view traditionally associated with Helmholtz [1867]; cf. Rock 1997). While every claim in this neck of the woods is controversial, it is plausible that perceptual processing operates according to different rules than cognition. For example, knowledge about the typical sizes of human beings does not prevent you from seeing the woman in Figure 15 as giant compared to the man on the other side of the room. And learning that the image is due to an illusory environment known as the Ames room does not make the effect go away.

Figure 15—Ames room[32]

Perceptual processing, one might conclude, is fundamentally different from cognitive processing and even encapsulated from cognition to some degree (e.g., Pylyshyn 1999; Firestone and Scholl 2016). Perhaps perceptual systems operate according to a proprietary "logic" (i.e., set of processing rules) that specify properties like size independently of the identity of represented objects and scenes. Perhaps, furthermore, perceptual systems cannot be penetrated by states activated in cognitive systems. It does not follow from these points, however, that the outputs of perceptual processing are nonconceptual.

Concepts are types of representations, and one and the same type of representation can in principle be deployable by a multiplicity of distinct processes. For example, you might form a mental image of the same particular configuration of visible properties on one occasion because

[32] 'Ames Room' by Ian Stannard, taken from https://www.flickr.com/photos/ silly_little_man/5132242358 on 9 August 2015 and altered to grayscale and size reduced. Licensed under Creative Commons License Attribution Share-Alike 2.0 Generic: https:// creativecommons.org/licenses/by-sa/2.0/

you read the words 'metal cylinder', on another occasion because you wanted a can of soda, and on another occasion because you saw the front of a car and mentally completed its hidden sides, including its exhaust pipe. The type identity of this representation is independent of its circumstances of deployment (and even the high-level and pragmatically relevant properties and individuals it functions to represent). Whether it is actually plausible that one would form the very same mental image via these various processes is immaterial; all that matters is that the situation, so described, is coherent. Concepts are paradigmatically deployed in cognitive processes like inference, but the idea that other processes might deploy concepts is not incoherent. There is no principled reason why a token conceptual *state* could not be the output of non-cognitive, proprietarily perceptual *processes*.

I therefore propose to examine the question of conceptualism independently of the question of inferentialism. My conclusion will be that PORs are conceptual outputs of perceptual processes. Conceptualism, though once popular, has largely fallen from favor (e.g., Crane 1992; Tye 1995; Smith 2002; Burge 2010a, 2010b; Hopp 2011). If my arguments for a mitigated conceptualism are sound, then the tendency to regard perception as wholly nonconceptual in respect of both the processing it involves and the states they operate on is a case of *under*-intellectualization. Though perceptual processing is plausibly distinct from cognitive processing, perceptual processing sometimes involves conceptual states. The tendency to under-intellectualize perception misses out on the fact that, while perceptual systems operate differently from cognitive systems, they sometimes deploy the very same types of representations.

As argued in Chapter 1 it is a *strength* of architectural accounts of the perception–cognition border that they allow perceptual systems to deliver conceptualized representations to cognition (see also Mandelbaum forthcoming). If perception and cognition utilize wholly separate representational formats—that is, if they speak different languages—then it is hard to see how cognition can immediately operate on perceptual outputs. If, however, perception delivers propositional structures composed of concepts, then cognitive systems can directly use those states as premises in inferences, map them to lexical representations, activate associated concepts, trigger memories and emotions, and so on. The widespread propensity to under-intellectualize perceptual states robs us of this sort of explanation of the relation between perception and the rest of the mind. Some measure of intellectualization therefore has *a priori* plausibility. But more importantly, as I will argue, both empirical evidence and philosophical reflection on the nature of concepts support conceptualism about PORs.

Many conditions for conceptuality have been proposed. Since the present project does not concern the *normative* or *epistemic* nature of concepts—that is, the distinctively normative role they play in the acquisition and transmission of justification or knowledge and in making a creature's mental states assessable for rationality—I will put aside normatively loaded conditions.[33] Instead, I focus on many of the leading conditions for conceptuality proposed in descriptive philosophy of cognitive science. These include systematicity, stimulus independence, amodality, inferential promiscuity, and composing into propositional structures. Before turning to each of these

---

[33] For the kind of normatively loaded views I have in mind, see, e.g., McDowell 1994; Brandom 1994. For now, I will simply beg the question against views on which thick kinds of normativity are constitutive of concepts, instead focusing on the question from a purely descriptive perspective.

conditions, however, it will prove useful to examine general issues about conceptuality and mental architecture.

§2—Concepts and Mental Architecture

There are two crucial (and related) choice points in determining conditions for conceptuality. First, is a representation conceptual in virtue of its *internal* or *relational* features? Second, is a representation conceptual in virtue of its *actual* or *counterfactual* functional role?

The internal/relational division might sound at odds with a functionalist mental ontology since, according to functionalism, mental kinds are individuated in virtue of some relational features, especially their relations to behavior and to other mental states (Lewis 1972, Fodor 1975). Nonetheless, we can distinguish between the functional role that is constitutive of a particular type of representation, and the role that it might acquire in virtue of some extrinsic factors. For instance, suppose some irrational person holds the belief that if 2+2=4, then the world will end tomorrow. In that unfortunate person's mind, the belief that 2+2=4 will have the functional role of making them think the world will end tomorrow. It does not seem that, from that fact alone, the functional role of making one think the world will end tomorrow is an internal feature of the belief that 2+2=4.[34] But on the other hand, it does seem that representing 4, and having a certain constituent structure, and perhaps many other features, *are* internal to the belief that 2+2=4. The internal

---

[34] On certain holistic views of cognition, this would indeed follow (e.g., Quine 1960). And on those views, it is not clear how an internal/relational distinction should be drawn. Indeed, it may be more accurate to say that such views simply reject the claim that there are any internal features of individual mental representations. That is arguably at the core of what is so problematic about such views; in any case, I put them aside here.

features of a representation are whatever features it has independently of how it might be used for a particular purpose at a particular time in a particular system, given other independent representations or architectural facts.

It is not always clear which aspects of a representation's occurrent causal influence in the mind are due to its internal features and which are artifacts of contingent intervening factors. While it is true that conceptual representations can play characteristic roles in inferences, for example, some representations will not actually be playing such a role at every particular time; they might not be activated at that time, or they might be activated but not participating in inferential transitions for some independent reason. It doesn't seem right that the representations cease to be conceptual in such cases. Performance errors can also prevent representations from fulfilling their characteristic roles. One such reason a representation might not fulfill its characteristic role is because of a subpersonal malfunction, such as neural damage; another is that resources might be depleted because of some other mental operation (e.g., cognitive load); another might be some sort of Freudian repression.

The architecture of a creature's mind can in principle have a significant impact on whether a conceptual representation exerts its potential causal powers. For example, perhaps central cognition is *fragmented* into different systems of beliefs such that relations between thoughts operate normally within a fragment but not across fragments (Lewis 1982, Egan 2008, Quilty-Dunn 2015, Mandelbaum 2016, Elga and Rayo ms). Fragmentation is typically posited to explain the persistence of contradictory beliefs. Given that human beings tend to eliminate contradictions between beliefs when those beliefs interact (e.g., in cases of cognitive dissonance—see Festinger

1957), contradictory beliefs that persist without resolution must be encapsulated from each other in some way. Fragmentation is supposed to provide diachronically flexible but synchronically rigid architectural barriers between "fragments," i.e., systems of beliefs and the concepts they comprise (Mandelbaum 2016). Some sort of fragmentation story is especially necessary if Spinozan theories of belief formation are true, according to which it is not psychologically possible for humans to entertain a proposition without believing it, in which case contradictory beliefs are ubiquitous (Gilbert 1991, Egan 2008, Mandelbaum 2014, Mandelbaum and Quilty-Dunn 2015). Fragmentation is a revision of the naïve Quinean view that one's beliefs are all interconnected in a coherent "web"; instead, fragmentation theorists propose an architecture consisting of disparate and internally coherent but mutually inconsistent micro-webs, or fragments.

Suppose that the thought P is in fragment *a* and the thought IF P THEN Q is in fragment *b*. Suppose, for the sake of argument, that typical inferential connections, such as modus ponens, are constitutive of the thought P. The thought P should normally interact with the thought IF P THEN Q to produce the thought Q; but, given the fragmentation of the two premise thoughts, the conclusion thought isn't formed. It does not seem that any representation's status as conceptual is thereby threatened. The representation is still the conceptual representation P in virtue of its internal features. It may also be conceptual partly in virtue of the truth of the following counterfactual: If the relevant thoughts were in the same fragment, then they would inferentially yield the thought Q.

One more example: Suppose that a certain cognitivist view of the psychological reality of syntactic rules is true, such that a given rule is part of an individual's I-language because that individual propositionally represents that rule in the language faculty (Fodor 1981). Though the

representation is quite isolated from conceptual representations in central cognition, it could still be true that they are concepts linked together in propositional structures. If this description is even conceptually *coherent*, then it is possible for a representation to be conceptual despite quite strict circumstances of architectural isolation.

The relation between conceptuality and mental architecture will be crucial in evaluating various proposed conditions for conceptuality, to which I turn now.

§3—The Varieties of Systematicity

3.1—Evans and the Generality Constraint. Evans (1982) famously argued that concepts obey a "Generality Constraint"—that is, thinkers who possess the singular concepts A and B and the predicate concepts F and G must be able to think the thoughts A IS F, A IS G, B IS F, and B IS G. The point of the Generality Constraint is that concepts are recombinable constituents of more complex concepts and propositional thoughts. Fodor and Pylyshyn (1988) emphasized what they called the *systematicity* of thought, which also asserts that thoughts involve systematically recombinable constituents. Though Evans had different purposes in mind than Fodor and Pylyshyn (see note 4, for example), I will treat systematicity and the Generality Constraint as picking out more or less the same feature of concepts.

Consider the most liberal possible form of the Generality Constraint:

PURE SYSTEMATICITY: A representation R belonging to a creature A must be able freely to combine at any time with any other concept in A's mind for R to be conceptual.[35]

---

[35] The formulations provided here abstract away from Evans's focus on "Ideas" (*de re* concepts of individuals) and predicate concepts and instead simply discusses concepts combining with each other in general. It might be best to

Almost everyone in the literature on concepts—including Evans—would argue that the kind of unconstrained and totally domain-general systematicity encoded in PURE SYSTEMATICITY is much too strong a condition for conceptuality.

The Generality Constraint functions in contemporary philosophy of mind in a manner akin to religious scripture: it is very widely accepted as true, often without argument, but there is a preponderance of inconsistent commentary on the primary text and vehement disagreement on how it should be interpreted and applied in particular cases.

As a footnote to his original statement of the Generality Constraint, Evans claimed that systematicity is constrained by whether the putative connection between the relevant concepts meets some semantic condition of appropriateness (1982, 101n17). Gauker provides an intuitive example: "A capacity to judge that John fell in the lake need not entail a capacity to judge that the lake fell in John" (2011, 4). Camp (2004) argues convincingly against this sort of semantic constraint. Since semantic appropriateness will not cut ice either way in what follows, however, I'll simply drop that qualification from here on out.

Carruthers (2009, 96ff) distinguishes between *weak* and *strong* systematicity. The weak systematicity requirement is that the relevant representation can (in the metaphysical, rather than logical or conceptual sense of 'can') combine with *some* other concepts that the creature could possess, whereas the strong requirement is that it can combine with *all* other concepts the creature

---

characterize things as he does and limit combinability conditions to singular concepts and predicate concepts, but nothing below hangs on this.

could possess. Carruthers argues for weak systematicity as the appropriate form of the Generality Constraint.

Peacocke (1992, 42–43) mentions that neural factors or Freudian repression can constrain systematicity. This point is much underappreciated and has significant consequences, as I'll argue below. As an example of Peacocke's point, consider the case of Kevin: it turns out that repression triggered by some traumatic childhood event causes Kevin's concept of his childhood nursery, which he does not consciously remember, to be isolated from some other concepts he possesses. Despite its repression, it does not seem right to say that the representation should thus fail to be conceptual. It is appropriate to talk about repressed *beliefs* and *desires*, and it is arguably a necessary truth that those states are conceptually structured, even if they are architecturally stored in a way that prevents them from exhibiting unrestricted combinatorial and inferential causal powers.

Reflecting on the possibility of repression, fragmentation, and other architectural divisions in central cognition illuminates a distinction between two kinds of systematicity. Systematicity is exhibited by a cognitive system and the particular battery of representations it contains; if two systems are encapsulated from each other, then systematicity will only be actualized within each system, not across systems. For instance, if the concept MORAL AGENT only figures in a fragment involved in moral judgment, and the concept ALPHA CENTAURI is stored in another fragment, then the thinker will be unable to think the thought ALPHA CENTAURI IS A MORAL AGENT. Call systematicity within a cognitive system with respect to the representations in that system *local* systematicity; call systematicity with respect to all representations in the mind, irrespective of fragmentation or encapsulation, *global* systematicity. The concept MORAL AGENT in this example

would actually exhibit local systematicity within its fragment, and would only counterfactually exhibit global systematicity; if the relevant fragments were merged, then the thinker could think ALPHA CENTAURI IS A MORAL AGENT.

Carruthers' weak–strong dichotomy is different from the local–global dichotomy I proposed above. His concern is with what combinations are metaphysically possible. He argues convincingly that the capacity to syntactically recombine some representations seems to be sufficient for the conceptuality of those representations, so the extra requirement that *every* conceptual representation that could be tokened in a given mind be so connectible to every other is too strong. Carruthers thus seeks to deny that it need be even metaphysically possible that all the concepts in a creature's mind could combine.

The global systematicity condition is not about what is metaphysically possible, but instead requires that the creature be in a position, holding its cognitive architecture fixed, to combine all its concepts. Even if we add in Evans's proviso that the requirement should be limited to categorial appropriateness, the global systematicity condition is still much too strong. The existence of fragmentation or other architectural divisions in central cognition, which would prevent the combination of representations in one fragment or system with those in another, together with the global systematicity requirement would jointly entail that adult human beings lack concepts. Since this conclusion is absurd, the global systematicity requirement should be rejected. Instead, systematicity should be localized within cognitive systems. We can still ask whether local systematicity should be weak or strong in Carruthers's sense, or whether strong systematicity should hold counterfactually across systems. For present purposes, however, the appropriate moral

149

to draw is that local systematicity is what matters in deciding whether discursive perceptual representations are conceptual.

One should accept the foregoing argument even if Freudian repression and belief fragmentation are not psychologically real. The point would still stand conceptually: the mind *might have been* fragmented, and yet there still would have been concepts in the various fragments. Similarly, if Freudian psychoanalysis is a flawed theory, it is not because it is metaphysically impossible for beliefs, desires, and other conceptually structured states to be repressed. Since we are only interested in the constitutive conditions of conceptuality, these mere possibilities suffice to establish the point at hand: we should not take functional properties that arise purely from such architectural constraints to be constitutive of concepts as such.

I propose that we shave the Generality Constraint down to the following principle:

> LOCAL SYSTEMATICITY: It must be metaphysically possible for a representation R in a subsystem S of creature A to combine with other concepts in S for R to be conceptual, whether or not R can combine with concepts elsewhere in A's mind.

Relativizing systematicity to subsystems allows LOCAL SYSTEMATICITY to express a deep truth about concepts—viz., that they compose with each other—while also recognizing the capacity of architectural borders to sharply constrain a concept's compositional scope.

Do discursive perceptual representations satisfy LOCAL SYSTEMATICITY? When tracking multiple objects, we can perceptually represent one object, O_1, as a white square labeled "A" in the top left of the visual field, and another, O_2, as a red circle labeled "B" in the bottom right. The empirical literature, as discussed in Chapter 2, shows that these sorts of representations figure in mid-level visual object perception (also see Quilty-Dunn 2016). It is hard to doubt that O_2, instead,

could be represented as a white square labeled "A", or a red square labeled "A", or a white circle labeled "B", etc. Not every concept the organism possesses is available to be used in object perception, so PORs do not display global systematicity. But they do appear to display local systematicity, in that any predicate applicable to visual objects can be applied to any visual object in various combinations.

3.2—Semantic systematicity and syntactic systematicity. One might object that even iconic representations exhibit systematicity in this sense. After all, as Mohan Matthen notes in criticizing Evans's Generality Constraint, "if one is able to entertain, in visual imagination for instance, images of a blue circle and of a red square, then one is able to entertain visual images of a blue square and a red circle" (2005, 80). This pushes us to yet another distinction: the difference between *semantic* systematicity and *syntactic* systematicity.

Semantic systematicity is the capacity of a representational system capable of representing multiple features together to represent different configurations of those features. Matthen's example is an instance of semantic systematicity. Syntactic systematicity is the capacity of a representational system to combine representational *vehicles* into larger syntactic structures. In an iconic representation of a blue square, since the representation has no constituent structure, it is not the case that there is one vehicle representing *blue* and another one representing *square*—one and the same syntactic part of a photograph of a blue square represents both the color and shape of the object. Since icons lack such constituent vehicles, icons cannot display systematicity of vehicles. Instead, they exhibit the capacity to coordinate contents in different ways without recombining distinct representational vehicles (see Chapter 4, ICONIC REPRESENTATION). This is precisely why

syntactic systematicity is a better guide to conceptuality than semantic systematicity; the former, but not the latter, requires separate vehicles to be recombined into larger structures, and this sort of syntactic recombinability is a hallmark of propositional thought.

I argued above that PORs are locally systematic, but that argument only sufficed to demonstrate semantic systematicity. To see whether PORs exhibit syntactic systematicity, we have to see whether PORs involve separate vehicles for each attributive element of representational content. It's not easy to determine an answer to this question. As noted above, however, we have independent reason to think that PORs are discursive. The semantic systematicity exhibited by iconic representations does not point toward syntactic systematicity, because icons lack constituents and are thus necessarily incapable of being syntactically systematic. But in the case of discursive representations, which have constituents, it is plausible to think that they represent distinct features via distinct types of constituents. Thus, showing semantic systematicity for discursive representations gives us good (defeasible) reason to think both that their constituents exhibit syntactic systematicity and also that the systematicity exhibited by those constituents explains the systematicity exhibited at the level of content.

Furthermore, while no positive evidence has been presented to think object files encode *any* features in an iconic format, there is good empirical reason to think object files encode at least *some* features in a discursive format. As mentioned in Chapter 2 (and Quilty-Dunn 2016), object files encode basic-level categories like *piano* and *fish* in an abstract, amodal format such that there is an object-specific preview benefit for *fish* even when the preview is the word fish and the probe is a picture of a fish (Gordon and Irwin 2000). It seems simply impossible for a representation to

encode information in that way without a discrete symbol (i.e., representational vehicle) for *fish*. In that case, not only is the argument-term in a POR a discursive label-like indexical constituent, but at least some of the predicate-terms are as well. In that case, given that those predicate-terms can obviously apply to other objects, there is very good empirical reason to think that PORs exhibit not only semantic systematicity but syntactic systematicity as well.

§4—Stimulus Independence

Elisabeth Camp (2009) argues that mere systematicity is insufficient for genuine concept possession, and that systematicity must be exhibited in a way that is independent of present stimulation for the systematically deployed representations to count as legitimate concepts. Systematicity-displaying representations that are beholden to stimulation are "so cognitively minimal that they encompass a type of engagement with the world that is more like passive triggering than active understanding" (Camp 2009, 303).[36] Stimulus-independent systematic representational capacities, on the other hand, are active in that the creature can deploy and combine concepts in a way that is not slavishly tied to what stimulation the creature is subjected to at a given moment. Since discursive perceptual representations are *perceptual*, and thus intimately tied to stimulation, perhaps they do not count as conceptual after all. Or perhaps, on the other hand, a representation's being conceptual isn't tied to its being *cognitive*, and conceptual representations can figure in perceptual as well as cognitive processes.

---

[36] I focus here on Camp's account rather than Burge's (2010b) account of "pure predication." I return to pure predication in the discussion of propositional structure below.

This debate gives off a faint odor of terminological quibbling. We need a taxonomy of different kinds of mental representations in accordance with different criteria of individuation; which ones we grant the honorific label 'conceptual' is not a serious theoretical issue unless it is tied to an explanatorily useful taxonomy. Furthermore, the claim that concepts are deployed in perception is compatible with the idea that such deployment is "more like passive triggering than active understanding." Indeed, if the perception–cognition border is architectural rather than representational, then we would expect that type-identical representations may be deployed both as a matter of passive triggering and as a matter of active understanding, depending on the context. On the other hand, it does seem of theoretical interest whether perception ordinarily involves conceptual representations. An approach that rules out the possibility by fiat should thus be put aside.

For debates about the role of concepts in perceptual systems to be substantive, we need to be able to distinguish the question of whether a representational vehicle is conceptual from the question of whether a mental process or system is cognitive. In answering the question of whether a vehicle used in some non-cognitive, perceptual system is conceptual, we should therefore ignore factors that are merely artifacts of the system's being perceptual.

This method of determining whether perceptual representations are conceptual is an instance of the following principle:

> ARCHITECTURAL INDEPENDENCE: If tokens of one type of representation, R_1, occur in cognitive system S_1 and tokens of another type, R_2, occur in system S_2, R_1 and R_2 are the same type of representation iff the difference in their functional roles can be explained entirely by the differences between S_1 and S_2.

ARCHITECTURAL INDEPENDENCE presupposes that representations can preserve type-identity across architectural divides. The mere possibility of belief fragmentation, again, suggests that representations can be conceptual even when they occur in different cognitive systems (recall the coherence, if not the empirical plausibility, of the Freudian repression thesis). Furthermore, it seems conceivable that two tokens of the thought that *p* might occur in two distinct fragments, even though their functional roles are quite different in the different fragments. So if the difference between two apparently distinct kinds of representations can be explained away simply by their occurring in distinct cognitive systems, then it seems fair to conclude that the representations are of the same kind.

One might grant that the difference is due to architectural factors but object that the two representations are *thereby* instances of distinct kinds, particularly if architectural factors bring along stark differences in functional role. In particular, the functional role of a representation in perception might be so different from that of paradigmatically conceptual representations in cognition that we might be tempted to deem the former representation nonconceptual even in cases where those functional differences are entirely explained by appeal to architectural differences. But in that case the debate at hand threatens to become merely verbal: one theorist decides to use the term 'conceptual' in a way that abstracts away from architectural divisions, and the other does not. To prevent the debate from becoming verbal, we should assume ARCHITECTURAL INDEPENDENCE and see whether the difference between discursive perceptual representations and those that can be deployed independently of stimulation is merely architectural or requires positing differences internal to the representations, abstracting away from architectural facts.

It does indeed seem to be a prominent feature of *cognition* that whether a representation is deployed is not determined solely by what stimulation the creature is currently experiencing. But that feature of cognition is due to facts about cognitive systems and the kinds of processes that figure in them, not necessarily due to any additional facts about the *representational vehicles* themselves. It seems possible that representations with all the same internal features as those that occur in stimulus-independent reasoning could occur instead in a system that only deploys representations given certain stimulation. One might balk at labeling those representations "conceptual" when they occur in stimulus-dependent systems, but the relevant difference seems to be entirely reducible to facts about cognitive architecture. In that case, the representations would be type-identical according to ARCHITECTURAL INDEPENDENCE.

Consider belief fragmentation in central cognition once again. Suppose that a certain fragment houses only implicit racist attitudes about a particular group, which attitudes are fragmented because they are incompatible with explicit egalitarian beliefs held by the same subject. Suppose that representations in the racist fragment are only deployed upon perceiving people of the relevant racial type. This possibility seems (a) entirely explained by architectural facts, and (b) entirely compatible with the representations' being conceptual. Indeed, Mandelbaum (2016) and Elga and Rayo (unpublished) argue that fragments are individuated by the contexts that elicit the accessing of their stored beliefs. This hypothesis supposes the beliefs stored in different fragments to be conceptual despite their activation being passively determined and thus "at the mercy of [thinkers'] environments" (Camp 2009, 290). Stimulus independence is thus not only explainable by appeal to architectural facts, but it also seems to be an unreliable criterion for distinguishing the

conceptual from the nonconceptual. And to reiterate a point from earlier, it doesn't much matter whether these fragmentation stories are *actually* true; it just matters that they describe coherent states of affairs in which representations maintain their status as conceptual despite being architecturally housed in such a way that they are only deployed in response to the satisfaction of certain elicitation conditions.

Stimulus independence is an architectural factor, and is best suited for theories that take the difference between perceptual representations and concepts to supervene entirely on where in the mind they occur. It makes perfect sense that a concept empiricist like Jesse Prinz (2002), for example, takes stimulus independence to be the basis of the distinction between nonconceptual perceptual representations and concepts. For Prinz, concepts simply are perceptual representations that are stored in long-term memory and become available for deployment outside their original perceptual contexts. In terms of the internal features of the representations, especially their format, there is no difference between concepts and perceptual representations on Prinz's theory. It is clear, therefore, why he employs a criterion for individuating perceptual representations from concepts that is purely architectural and allows perceptual and conceptual representations to be internally alike. The same motivation for stimulus independence might hold for a sort of pure language-of-thought view (Pylyshyn 2003 may be an example, particularly given his apparent opposition to positing iconic representations at any stage of perceptual processing), according to which all representations in central cognition and perceptual systems are amodal discursive representations.

But for those who reject concept empiricism and pure language-of-thought views, we must expect there to be *internal* differences between nonconceptual perceptual representations and

concepts—differences that do not just reduce to facts about where in the mind they occur. Stimulus independence may be a good way of distinguishing perceptual systems from cognitive systems (though it may not do justice to the evidence in favor of central modules and belief fragmentation). But that architectural division does not necessarily map on perfectly to the distinction between nonconceptual and conceptual representations, particularly if we reject views like those discussed in the previous paragraph. Distinguishing nonconceptual perceptual representations from concepts by appeal to differences in format is far more congenial to mainstream non-empiricist accounts of concepts than doing so by appeal to stimulus independence.

Recall, furthermore, the dialectic between representational and architectural approaches to the perception-cognition border. The project at the outset of this chapter was to see whether the representational strategy can be resisted by arguing that the discursive outputs of perceptual systems are type identical to concepts deployed in cognitive systems. The foregoing discussion suggests that stimulus independence is an irrelevant criterion for that project, since stimulus independence is consistent with concepts and percepts being internally type-identical (as Prinz would argue). Stimulus independence is thus not an adequate condition to appeal to in evaluating the representational strategy for distinguishing perceptual from cognitive systems.

Finally, Gordon and Irwin (2000) showed that object files incorporate symbols standing for abstract properties like *fish* that are primed by words (e.g., 'fish') and tested by pictures (e.g., a picture of a fish). Are these symbols type-identical to the concept FISH? Note how unhelpful the criterion of stimulus independence is for answering this question. PORs are stimulus dependent token representations, which are only deployed by the visual system via non-cognitive perceptual

processes in response to stimulation. It follows that the deployment of the *token* symbols in a POR is stimulus dependent. The question at hand, however, concerns whether those token symbols are of the *type* FISH. Assuming stimulus independence as a criterion on conceptuality, the question becomes whether the symbol tokened in perceptual contexts is the same type of symbol tokened in non-perceptual, paradigmatically cognitive contexts (as in, for example, when one wonders what kind of fish live at the bottom of the Atlantic). In this case, however, stimulus independence is unilluminating. We need to have independent criteria on conceptuality in order to answer the question of whether token representations that figure in PORs are type identical to token concepts that figure in thought. Even if we accept stimulus independence as a criterion on a (type of) representation's being conceptual, it is unhelpful in determining whether conceptualism about perception is true.

§5—Amodality

One might object that concepts are constitutively amodal (Fodor 1975), and that PORs are modality specific and so nonconceptual. The supposed amodality of concepts is controversial (Prinz 2002), but we can assume for the sake of argument that concepts are constitutively amodal only to make the case against the conceptuality of object representations even stronger. PORs, however, do not obviously fail to meet this criterion.

There are two ways in which a representation can be modality specific. A representation is *extrinsically* visual if it is visual only insofar as it is deployed in the visual system. A representation is *intrinsically* visual if it is visual in a way that is determined by its format, independently of the

conditions in which it is deployed. An example of a representation of the second kind might be a visual iconic representation. Visual iconic representations not only exhibit neural retinotopy (Frisby and Stone 2010; see Chapter 2 of this dissertation), but they also coordinate color with shape and location in a way that no representations in other modalities do (see Chapter 4 of this dissertation). Visual icons lack separate constituents for shape and location but instead have selfsame iconic parts that determine locations along quality dimensions of both color and location; arguably, visual iconic representations are intrinsically *visual* because they coordinate color with location/shape in this fashion, while auditory iconic representations are intrinsically *auditory* because they coordinate pitch, volume, etc., with location in a similar fashion. Moving outside the head, a painting is intrinsically not an auditory representation, but a visual one, and this is arguably due in part to the way it coordinates color with shape and location without contingent syntactic concatenations of representations of those features. Since discursive perceptual representations are not iconic, they do not exhibit this kind of intrinsic modality specificity.

Perhaps PORs, due to their architectural position in mid-level vision, are *ipso facto* (extrinsically) modality specific. If so, their apparent nonconceptuality might reduce to architectural factors. It's not clear in that case that *extrinsic* amodality should be taken to be constitutive of conceptuality. As argued in the discussion of stimulus independence above, mere architectural isolation of a representation is compatible with its being type-identical to conceptual representations in central cognitive systems, and thus should not be regarded as a sufficient condition for that representation's being nonconceptual. The same point applies here with respect to extrinsic modality specificity.

It is far from clear, however, that representations in mid-level vision or later in visual processing are really encapsulated within vision. Indeed, Spelke (1988) argues that the capacity to represent objects is independent of a particular modality on grounds that object representations exhibit systematic crossmodal effects (cf. Scholl and Leslie 1999). Carey (2009) also argues that object representations are not distinctly visual because they persist despite occlusion. Scholl and Leslie (1999) argue against claims by developmental psychologists like Spelke that object representation is totally non-perceptual and "maximally central" (26). Instead, they defend what they call the "object-indexing framework" where the representations responsible for Spelke's (and others') results are the same as those responsible for multiple-object tracking and object files, i.e., mid-level visual object representation. But even Scholl and Leslie articulate the apparent lack of modality specificity of object representation:

> An interesting challenge to the object-indexing framework comes from cross-modal aspects of the object concept[.]…Some contact-mechanical aspects of the object concept, at least, do seem to be cross-modally sensitive…and the same may hold for spatiotemporal aspects[.]…The apparent conflict lies in the fact that mechanisms of object-based attention have primarily been proposed to function over only a *visual* domain and to be somewhat distinct from other modalities: it has often been noted that visual and auditory processing, for instance, share several surprisingly similar mechanisms[.]…In addition, although it is the visual aspects of Pylyshyn's visual-indexing framework that have received the most experimental attention, visual indexes are primarily intended to be part of a visuo-*spatial* system, and Pylyshyn's model explicitly includes a proprioceptive component[.]…It is possible that object indexing operates in a single cross-modal space, but this is not a necessary aspect of our model.
>
> (Scholl & Leslie 1999, 59; citations omitted)

More recently, Jordan et al. have shown that the object-specific preview benefit (Kahneman et al. 1992) works cross-modally—specifically, object files formed through vision can be accessed by audition. The initial preview information presented with a given object was a picture of a familiar

object, such as a piano, and the target information, which participants had to confirm or disconfirm was a "match" with the preview information, was a corresponding sound, such as the sound of a struck piano key. Subjects exhibited an object-specific preview benefit for sounds that matched the visual preview information, even when resources for verbal rehearsal were depleted by a load task in which they had to silently verbally repeat a string of numbers in inner speech. They conclude that "object files can operate *across* visual and auditory modalities" (2010, 501; emphasis theirs).

PORs enable access to encoded representations from multiple sensory modalities, encoding kinds like *piano* in a way that is not limited to the visual appearance or sound of pianos, which belies the claim that they are distinctly and intrinsically visual. PORs are not modality specific in any substantive sense: they are not iconic; as mentioned above, they encode non-vision-specific properties of objects; and despite functioning in visual processing and being commonly referred to as constructions of "mid-level vision," they do not appear to be architecturally limited to vision alone. Jordan et al. conjecture that object files "store object-related information in an amodal format that can be flexibly accessed across senses" (2010, 500). It appears that PORs, like conceptual representations, are not only discursive but intrinsically amodal.

§6—Inferential Promiscuity

A representation is inferentially promiscuous if it can figure as a premise in an inferential transition (Stich 1983).[37] Since inferences are a paradigmatic case of a process that operates on conceptual

---

[37] One might reserve the term 'premise' for a proposition that is logically related to the other elements of an argument rather than a mental representation, in which case the text should say "premise-representation" or something similar instead of simply "premise." I gloss over this fine point in the text.

representations, inferential promiscuity is a promising criterion for conceptuality (see, e.g., Peacocke 1992).

Carey (2009) and others (e.g., Scholl and Leslie 1999) argue that PORs in adults (i.e., visual indexes and the object files they are linked to) involve the same underlying capacity to represent objects that is studied in developmental psychology (e.g., Spelke 1988). Carey (2009) and Spelke (1988) argue based on the developmental literature that infants use object representations as premises in inferences. For example, a POR might represent an object as solid and cause the infant to infer on that basis that the object will not go through a solid shelf when dropped, thus generating surprise (as measured by an increase in looking time) when the object is revealed to be below the shelf (Spelke et al. 1992; see also Carey 106ff). Perhaps, in that case, PORs are indeed inferentially promiscuous.

Despite being intuitive, however, this form of explanation is not mandatory. One could reasonably object that the computations responsible for looking-time effects like these do not constitute genuine inference. This example illustrates how hard it is to show that a given representation is inferentially promiscuous in a non-question-begging way. Every time we might point to an apparently inferential transition that the representation initiates, an opponent could reply that the transition isn't *really* inferential. For instance, Helmholtz (1867) took all perceptual processing to involve inference, but today many theorists regard much of perceptual processing as non-inferential computation (e.g., Burge 2010a). Merely showing that a process transforms one representation into another and generates behavioral responses in a lawlike way is not sufficient to

show that the process is inferential. To better understand inferential promiscuity as a criterion on conceptuality, therefore, we need to understand what constitutes an inferential process.

First, however, it will help to get clearer about the modality of 'can' in characterizing an inferentially promiscuous representation as one that "can" figure in inference. Does this criterion require occurrent inferential promiscuity such that the concept-holder presently has the ability to use it in inference? Or does it require merely that the representation is the *type* of representation that could figure in inferences, even if it never actually does?

Some philosophers have argued that concepts simply *are* inferential roles (e.g., Peacocke 1992). In that case, inference is metaphysically prior to conceptuality. It follows that a representation cannot be a concept unless the creature who possesses it also possesses the occurrent capacity to use it in inference. Inferential-role accounts of conceptuality therefore seem to entail that *occurrent* inferential promiscuity is a necessary condition for a representation's being conceptual.

Fodor (2004) argues that inferential-role accounts of conceptuality are left without a substantive theory of what inference is. Moreover, I argued in Chapter 1 that at least a superficial (in Schwitzgebel's [2013] sense) form of dispositionalism that casts representations in terms of occurrent dispositions for processes to take place is not feasible. A computational account of mental processes construes such processes as causal operations on mental particulars and thereby enables a substantive causal-mechanical explanation of behavior. Inferential-role accounts cannot construe inferences as operations on mental particulars, since concepts, on such accounts, simply are capacities to engage in certain inferences. Since a capacity to engage in an inference is not a mental

particular, it is *a fortiori* not a mental particular on which some computational process can operate (cf. Peacocke 2004). Thus computational theories of mind, even those that otherwise bear little resemblance to traditional language-of-thought theories (such as Prinz 2002), are united in their commitment to concepts' being mental particulars.

If inferential-role views of conceptuality are false, then a representation might not need to be occurrently disposed to figure in inference to be conceptual. Instead it might be that concepts are symbols that are constitutively *apt* to figure in inference. There is some independent reason to think this is the case. Some beliefs might be quarantined via repression or fragmentation, for example, such that, even though *qua* beliefs they are apt to function as premises in inferences, the architecture prevents this capacity from being exercised. Here as elsewhere it seems that the important constitutive question is whether a representation is apt to perform a certain function rather than whether it happens to be situated in a mind that enables it to do so. Indeed, if computational approaches to inference are correct, then understanding the nature of inference may require understanding the structures of conceptual states rather than vice versa. For computationalism assumes that inference, like other computational mental processes, is a certain kind of operation on certain kinds of symbols. In that case, knowing which kinds of symbols figure in inference and how their structures enable them to play the roles that they do is necessary for an adequate understanding of inference. Characterizing inferential transitions, on this approach, requires first characterizing the states involved in those transitions.

Computationalist views thus invite a "representation-first" approach to concepts, while inferential-role views invite an "inference-first" approach. Note that these are approaches to the

metaphysical question of which properties are constitutive of conceptuality; the *epistemological* question of what constitutes good evidence for a state's being conceptual is a separate one. One can adopt a representation-first approach to concepts while also taking the apparent presence of inferential capacities to be very good evidence in favor of judging a representation to be conceptual. And one can adopt an inference-first approach while also taking the apparent presence of an amodal discursive format, for example, to be very good evidence in favor of judging a representation to be conceptual (though it is not as clear how to understand format on an inference-first approach).

When considering the metaphysical question of conceptuality from a computationalist perspective, we should first consider the kinds of states that are constitutively apt to function as premises in inferential transitions. From a computationalist perspective, the leading candidates for the states involved in paradigmatic conceptual inference are *propositionally structured* mental representations (e.g., Fodor 1975; 1987; Fodor & Pylyshyn 1988; Pylyshyn 1984; Braine & O'Brien 1998).[38] Inferential transitions, on a promising version of this view, are transitions between propositionally structured mental representations in virtue of their constituent structure alone, in accordance with some built-in formal rule of mental logic such as modus ponens (Quilty-Dunn and Mandelbaum unpublished b). Since PORs are discursive and thus have constituent structures,

---

[38] Camp (2007) and Rescorla (2009a) argue that creatures can think and draw inferences by means of maplike representations. Like them, I assume that at least much of paradigmatically cognitive inference in humans involves non-maplike, propositionally structured representations figuring in logical inference. I reserve the term 'concept' for constituents of such propositional structures. One might insist that the constituents of maplike representations in cognition count as conceptual. Dialectically, however, appealing to this looser notion of conceptuality would seem to be unfair to anti-conceptualists about perception. I therefore assume the stricter notion of conceptuality that demands that concepts be apt to compose into propositional structures.

then they may be constitutively apt to function as premises in structure-sensitive inferential transitions. The question turns on the specific sort of discursive format instantiated by PORs.

To sum up: inferential promiscuity requires propositional structure. It is because thoughts of the form A IS F have the propositional structure they do that they combine with thoughts of the form IF A IS F THEN A IS G to inferentially yield the conclusion A IS G. We can therefore make headway on the question of whether PORs are apt to function as premises in inferential transitions by seeing whether they exhibit propositional structure.

§7—Propositional Structure

The format of PORs is discursive, but it does not necessarily follow that it is propositional. For instance, maps have a discursive format, but are not syntactically propositional (Camp 2007; Rescorla 2009a) and also contain non-discursive, iconic elements (Kulvicki 2015b). One can imagine taking an aerial photograph of a public park and putting discursive symbols on it (e.g., a 'B' to mark a bathroom, and 'F' to mark a water fountain) and thereby composing a map. And maps differ in format from graphs, blueprints, hand gestures, and other types of representations that may be discursive without being propositional and wholly non-iconic.

Propositional structure constitutively involves a kind of relation between representational vehicles that mirrors a predicative relation between their contents, thereby enabling the complex vehicle to express predication. As Camp puts it, "some sort of functional relation among syntactic constituents maps onto some sort of logical or metaphysical relation among the semantic values of

those constituents" (2007, 157).[39] We can call this feature of propositional representations *syntactic predication*, since it is directly analogous to, and functions to express, semantic predication (see Chapter 4 for further discussion of syntactic and semantic predication).

The mere attribution of a property to an individual is not sufficient for syntactic predication, since property attribution can be accomplished without complex syntactic structure. A picture of a red barn, for example, may attribute redness to the barn, but it does not possess separate *vehicles* for the barn and its color—one cannot cut out the part that represents the barn but leave the part that represents its redness, or vice versa. Moreover, some philosophers might separate mere attribution of properties from semantic predication and argue that iconic representations accomplish the former but not the latter (e.g., Burge 2010a), in which case icons *a fortiori* fail to instantiate syntactic predication.

Even representations that have semantically predicative propositional contents do not necessarily instantiate syntactic predication. For example, suppose one quail hunter yells "Quail!" to another. This utterance arguably has a propositional content like <There is at least one quail nearby> or <That is a quail>, but it is syntactically atomic; the semantic predicative relation is in no way made explicit in the syntax of the linguistic token (cf. Quine 1960, Chapter 2).

---

[39] Camp's formulation seems more favorable to Russellian rather than Fregean accounts of propositions. Her formulation takes the distinctly proposition-like structure of propositional representations to consist in a mapping of syntactic structure onto logical/metaphysical relations among the extensions of their constituents. This makes the most sense if one takes the constituents of propositions to be those extensions. If the constituents of propositions are Fregean senses, on the other hand, then the mapping relation that marks a representation as propositional should be between its syntactic structure and the semantic relations among the constituent senses of the Fregean proposition. I will describe the mapping relation as between the way constituents are syntactically ordered and the way their contents are semantically ordered, intending thereby to remain as neutral as possible between Fregean and Russellian accounts of propositions. I put unstructured theories of propositions aside for simplicity's sake, though I take them up again in Chapter 4.

Another illustrative example comes from maps and the markers they contain. Maps involve iconic representations of the spatial layout of the represented terrain. They also, however, involve markers that function non-iconically. Markers are discrete symbols that are deployed to predicate the instantiation of some property at a location corresponding to their location on the map (such as the 'B' to predicate the presence of a bathroom at some location in the example above).[40] Even if, as a matter of fact, the marker is an icon—e.g., a picture of a house that marks the presence of a house—it functions as a discrete symbol relative to the rest of the map. To predicate the presence of treasure at a location on a map, one might write the sentence, 'The treasure is here.' 'The treasure is here' exhibits both semantic and syntactic predication, since there are separate vehicles for the property and the location represented, and there is a syntactic relation between the two that mirrors the semantic relation between their contents. Instead, however, one might simply draw an "X". The "X" predicates the presence of treasure at the corresponding location in the terrain, and thereby has the propositional content <there is treasure here>. The "X" does not, however, express that content in a syntactically explicit fashion. It is syntactically atomic and lacks constituents that compose in a fashion that recapitulates the semantic relation of predication present in its propositional content. The fact that a representation accomplishes semantic predication, therefore, does not entail that the representation exhibits syntactic predication.

With these caveats in mind, we can now turn to the question of whether PORs exhibit syntactic predication. The distinction between the indexical constituents of PORs (i.e., visual

---

[40] Rescorla (2009a; 2009b) offers interesting arguments for the conclusion that maps do not accomplish predication, which I consider critically in Chapter 4.

indexes, or "FINSTs") and feature constituents requires that there be separate vehicles corresponding to individual objects and the properties attributed to them by the POR. Furthermore, PORs can be accurate or inaccurate depending on whether the objects they pick out have the properties they represent. For example, a POR that stands for object O_1 and encodes the property *red* is accurate iff O_1 is red, and thereby has the content $<O_1$ is red$>$. In that case, their evaluability for accuracy arises from the way their constituent feature representations are connected to their constituent indexical representations. That is, PORs have separate vehicles for represented individuals and properties, and those vehicles are syntactically ordered in a way that mirrors the semantic relation between their constituents. PORs, therefore, exhibit syntactic predication. PORs are propositionally structured perceptual representations.

According to Burge (2010b), a representation must exhibit "pure" predication to count as genuinely propositional. Pure predication consists in the applicability of a predicate outside of purely referential contexts. For example, the application of the predicate 'man' in the phrase 'that man is tall' is not pure, since it appears within the scope of the demonstrative referential noun phrase. The application of 'tall', on the other hand, is an instance of pure predication. Tallness is attributed to the referent of 'that man', but outside the scope of the referential noun phrase. Burge further claims that, since the content of perception is always akin to a referential noun phrase, perception never exhibits pure predication.

It is unclear whether Burge should be read as characterizing the structure of the vehicles of representation rather than the structure of content. After all, he claims that the referential-noun-phrase-like structure of <u>That F</u> is present not only in perception, but also in photographs (2005, 6)

and maps (2010a, 540). Photographs and maps do not seem to have a THAT F *syntactic* structure—for example, what would be the demonstrative syntactic constituent of a photograph? If Burge is indeed not describing the syntactic apparatus necessary for propositional structure, then his view may be compatible with the thesis that PORs have a propositional structure. He may instead argue that they lack a certain kind of propositional *content* that is distinctive of at least some forms of rational thought. This claim requires the decidedly Fregean assumption that there is more than one kind of content including some that are non-propositional, which I cannot address here. If Burge's view just requires that propositional content involves pure predication, then it is consistent with the claim that there might be representations with a propositional syntactic structure that nonetheless lack propositional content. Since my concern is the structure of PORs rather than the nature of their content, this view would for present purposes be benign.

It may nonetheless be fruitful to take Burge's condition of pure predication as a proposal about the syntactic structure of propositional thought. One way of understanding pure predication is as requiring stimulus independence. That is, for a feature representation to function in a purely predicative fashion, it must be capable of being applied outside of perceptual contexts. This seems to follow from Burge's claim that pure predication requires application outside the scope of a referential phrase together with his claim that perception always has a referential structure. But in that case, the pure-predication criterion faces all the same problems as the stimulus-independence criterion. Furthermore, the question of whether perceptual processes ever deliver propositionally structured representations (such as PORs) seems like a thoroughly empirical one. But Burge's

criterion of pure predication, like all stimulus-independence-based views of conceptuality, rules out the possibility *a priori*.

The pure predication criterion seems to be ambiguous between two claims. One is that, for a representation to be purely predicatively applied, the *representation* must be such that it could be deployed in non-referential contexts. The other claim is that the *creature whose representation it is* must be capable of deploying that representation in non-referential contexts. It seems possible (i) that a creature might form representations with a structure that would in principle allow those representations to be taken apart and deployed separately in other contexts, but also (ii) that the creature lacks the cognitive apparatus to take apart and redeploy those representations. Pure predication thus seems not to be a feature of the format of a representation, but rather to be a feature of *cognitive systems*. Burge's invocation of pure predication as a condition on propositional structure thus seems to focus on abilities of cognitive systems rather than representational format. But propositional structure is better understood as a matter of representational format.

It is precisely because a cognitive system operates on representations ordered into complex propositional structures that it is capable of pulling those simple representations apart and redeploying them elsewhere. Propositional format is ontologically prior to cognitive processing, and the pure predication criterion reverses that order. The redeployability of a representation outside purely referential, perceptual contexts may be good *evidence* of propositional format, but it should not be taken to be constitutive of propositional format. At best, doing so requires that a creature have amodal cognitive capacities in order to have propositional representations and thus deems the thesis that non-cognitive states can be propositional false by fiat. At worst, it robs us of

causal-mechanical explanations of behavior that construe cognitive processes as computational operations on (ontologically prior) propositionally structured representations.

I have been criticizing the use of pure predication as a condition on propositional structure and arguing that PORs might be propositional even if they fail to exhibit pure predication. It is consistent with that criticism, however, that PORs also happen to exhibit pure predication. I argued above that PORs are intrinsically amodal. They represent kinds like *fish* and *hammer* in a way that is totally untethered to low-level features. Moreover, these representations are mapped to lexical items like the word 'fish' and to auditory representations like the sound of a hammer banging. This provides independent reason for thinking these representations are the same as those we use when thinking about fish or hammers—that is, that they are the concepts FISH and HAMMER. The fact that they are deployed perceptually does not rule out this possibility, even assuming pure predication as a condition on propositional structure. If we were to try to find evidence of a purely predicative representation applied in visual perception, we would try to find evidence of its availability for linguistic and non-visual processes. It is plausible therefore that at least some of the constituents of PORs function outside purely referential, perceptual contexts and are therefore applied in a purely predicative fashion.

§8—Conclusion

An empirical hypothesis for object perception emerges from this discussion: Object representations in perceptual systems are compartmentalized conceptual representations. They exhibit some of the standard features of concepts, such as discursivity and constituent structure, segmenting stimuli

into individuals, representing/categorizing, local systematicity, amodality, propositional structure, and inferential promiscuity. They don't occurrently exhibit others, such as global systematicity, because of their architectural position outside central cognition, but they may possess them in a counterfactual sense such that they would occurrently exhibit them if placed in central cognition.

It could instead be true, as Scholl and Leslie write, that the "mechanisms and processes which drive infants' discriminative abilities may best be characterized as neither 'perceptual' nor 'conceptual,' but somewhere in between" (1999, 27; cf. Shea 2014). There is a substantive sense in which they are correct. To be sure, PORs are not *cognitive* representations; they are not the result of operations in central cognition. Indeed, I argued in Chapter 2 that the evidence overwhelmingly suggests that PORs are perceptual. But I have given some reason to think that they are nonetheless a species of conceptual representation. There is an apparent difference between such representations and fully promiscuous conceptual representations in central cognition, but this need not arise from a difference in kind. In the absence of some more substantive proposal that accounts for both the iconicity of early vision and the discursivity of PORs, I take it as an intact working hypothesis that such representations are conceptual despite their occurrence in mid-level perceptual systems rather than central cognition.

We can thus draw a distinction between four kinds of representations: (i) cognitive conceptual representations, which are discursive representations in central cognition such as ordinary beliefs; (ii) perceptual conceptual representations, which are discursive representations in perceptual systems such as PORs; (iii) cognitive nonconceptual representations, which are nondiscursive representations in central cognition such as certain forms of imagery and episodic

memory; and (iv) perceptual nonconceptual representations, which are nondiscursive representations in perception such as iconic perceptual representations.

One might object to this whole line of argument on grounds that being conceptual entails being cognitive, and being cognitive entails not being perceptual (e.g., Raftopoulous 2014).[41] *Of course*, one might say, *being conceptual is an architectural matter. Concepts live in central cognition and nowhere else—that's part of what the term 'conceptual' means. So we know* a priori *that nothing in perception* could *be conceptual.* This perspective is perhaps legitimate, but as noted above, the dispute at this point seems to devolve into a futile war between terminological preferences. It does seem, on the other hand, that there is a substantive theoretical question whether perception is conceptual or nonconceptual.

Insofar as this question is substantive and there is an architectural division between perception and cognition, I have argued, any attempt to answer it must abstract away from merely architectural differences and accept ARCHITECTURAL INDEPENDENCE. Since conceptual representations in central cognition and some perceptual representations are internally alike, and

---

[41] Raftopoulos (2014, 614–615) gives additional arguments for why concepts can't be encapsulated, but they are tailor-made for early vision and so don't successfully undermine the arguments in this chapter. He argues (i) that a certain account from Pylyshyn doesn't posit genuine conceptual representations but rather something implicit, (ii) that representations in early vision are iconic and so don't logically combine or figure in inferences, (iii) that early visual representations don't allow for reidentification, and (iv) they don't satisfy the Generality Constraint. Claims (i) and (ii) aren't relevant to the proposal at hand, claim (iii) doesn't seem to apply to object representations in mid-level vision nor to face perception or other forms of late vision, and claim (iv) is argued against at length above for representations in mid-level vision and up. Raftopoulos may in fact agree with the thesis of this chapter, since he thinks only early vision is nonconceptual and iconic. But the arguments just mentioned seek to show that encapsulation is sufficient for nonconceptuality, and they do not succeed at establishing that conclusion. Moreover, object perception is synchronically encapsulated despite deploying concepts.

their differences are entirely explicable by appeal to architectural divisions, the answer to the *substantive* question whether PORs are conceptual seems to be yes.

ICONIC REPRESENTATION

Chapter 4

"Does my visual image of this tree, of this chair, consist of parts? And what are its simple component parts?"

—Ludwig Wittgenstein, Philosophical Investigations*, §47*

§1. Problematizing Icons

Early perceptual systems and mental imagery involve modality-specific iconic mental representations—perceptual icons, for short. Like photographs or paintings, perceptual icons are profoundly un-language-like. In particular, every part of an icon corresponds to a part of the scene represented by the whole icon; icons lack canonical decompositions into privileged constituents, rendering them fundamentally different from anything resembling human language or propositional thought (Kosslyn 1980, 1994; Johnson-Laird 2006; Fodor 2007; Carey 2009; Kulvicki 2015a; Quilty-Dunn 2016).[42] As I argued in Chapter 2, these syntactic facts allow icons to figure in substantive explanations of high-capacity short-term perceptual memory capacities (e.g., Sperling 1960; Coltheart 1980; Bradley and Pearson 2012), as well as mental imagery (Kosslyn 1980, 1994). I also argued that icons may help elucidate hallmarks of perception such as representation *via*

---

[42] Though perhaps not so different from so-called "analog magnitude" representations used by animals and infants to compute numerical magnitudes (Carey 2009, Chapter 4; Mandelbaum 2013).

spatial frameworks, phenomenal richness and fine-grainedness, and the cognitive impenetrability of early perceptual systems.

The fact that iconic representation is so unlike linguistic representation also presents several special problems for icons that do not arise for language-like representations such as propositional thoughts (at least as understood in terms of the classical computational theory of mind—see Fodor and Pylyshyn 1988). I'll begin by detailing these problems and outlining how representations with constituent structure easily solve them. I'll ultimately argue that icons solve all of these problems by coordinating analog values together, without adverting to constituent structure. My account is intended to characterize not only perceptual icons, but also the bare-bones syntactic and semantic properties of icons more generally, including paintings, graphs, photographs and emojis. The resulting account is meant to provide a scaffold for understanding iconic representation both inside and outside the mind.

1.1—The problem of local processing. The mind is probably computational. That is, (at least some) mental processes should be understood as rule-governed operations on mental representations. Furthermore, computation is probably syntactically driven. That is, computational processes operate on locally individuated symbols, irrespective of the meanings of those symbols. As mentioned in Chapter 1, the contents of mental states tend to be *wide*, in that they depend on relations to the environment (and, typically, to their referents). The concept CICERO has the content it does not (solely) in virtue of anything internal to the minds of those who possess it, but rather (at least partly) in virtue of causal chains ultimately terminating in the man himself (Kripke 1980; cf. Burge 1979). These causal chains are completely extrinsic to cognitive systems involved in

reasoning about Cicero, so these systems must operate on *local* properties of the concept.[43] I have

framed this point in terms of the *syntax* of the relevant mental representations, in contrast with

their *semantics*: syntactic properties are the locally individuated, content-independent but

semantically exploitable properties of symbols, whereas semantic properties are those constitutive

of sense, reference, evaluability for accuracy, and so on.[44]

According to the classical computationalist story—canonically defended by Fodor and

Pylyshyn (1988)—thoughts have constituent structures articulated in an innate, amodal language

of thought, and the rules built into central cognition specify types of constituent structure rather

than the (arguably inaccessible) semantic properties of thoughts or their conceptual constituents

(see also Fodor 1975, 1987; Pylyshyn 1984). Different elements of this story have been challenged

by computationalists. Some argue that some thoughts are map-like (Camp 2007; Rescorla 2009a),

others argue that the constituents of thoughts are modality specific and acquired through

perception (Prinz 2002), others argue that thought is couched in natural language (Gauker 2011),

---

[43] 'Local' here means internal to a representational vehicle *embedded within a combinatorial system.* The orthography of some atomic linguistic item might suffice to distinguish it from some others of the same syntactic category (e.g., 'dog' and 'cat'), but won't determine whether it is a noun phrase or a predicate, for example. It's only relative to the combinatorial syntax of the language that some string of letters is a member of this or that syntactic category. Thus the syntactic properties of a symbol will be determined by its internal (e.g., orthographic) features and the way that it combines with other symbols within a system.

[44] Rescorla (2014) argues that computational transitions can be sensitive to semantics, but his arguments involve interventionist counterfactuals in which a computational process cannot be altered without altering locally individuated or "indigenous" semantic properties. These indigenous semantic properties are said to arise from embedding computational processes such that the output of one functions as input to another; in such cases, altering the content of a symbol necessitates altering its computational role, and computational operations on it are thus sensitive to its semantic properties. These interventionist counterfactuals might be true in certain cases of embedded computational systems, and thus some forms of computation might be "sensitive" to semantics (especially on an interventionist model of causation). This point does not undermine the claim that computational processes operate on syntactic features of mental representations. Indeed, Rescorla's arguments may be interpretable as concluding that computational sensitivity to semantics *emerges from* operations on nonsemantically individuated symbols (though Rescorla himself might resist this interpretation).

and still others accept much of the classical framework but offer anti-nativist accounts of concept acquisition (Carey 2009). These alternate formulations, however, tend to presuppose the core computationalist picture: thoughts have constituent structures, which allows cognitive systems to instantiate mental processes like inference by enacting computational operations on those structures via purely syntactic rules (see Braine and O'Brien 1998; Quilty-Dunn and Mandelbaum unpublished b).

The problem of local processing may by this point be clear. Mainstream theories of mental processing for decades have understood mental processes to be computational, and have understood computational mental processes as operations on constituent structures. Icons lack constituent structures, so they cannot figure in computational processes as ordinarily understood. But perceptual icons surely are inputs to and outputs of mental processes: icons in early perception are produced on the basis of retinal input and feed into later perceptual processes, and may even feed directly to conceptualization (as plausibly occurs in partial report paradigms—Sperling 1960) and motor systems; icons in mental imagery are generated on the basis of cognitive intentions and information stored in long-term memory, are transformed within imagery systems in operations like mental rotation (Shepard and Metzler 1971) and scanning (Kosslyn et al. 1978), and enable the acquisition of beliefs in central cognition (such as beliefs about spatial relations between imaged objects—Finke and Pinker 1982).

The same arguments that motivate computationalism about mental processes generally, however, should also motivate a computationalist understanding of how perceptual icons figure in mental processes such as those described above. One might marry the notion of computation to

constituent structure such that processing icons could not count as computation; but that move would define the problem away rather than solving it. We are still left with the problem of how icons figure in local processing (i.e., operations on locally individuated properties of icons), whether or not such processing counts as Computation-with-a-capital-C. The classical computationalist picture of mental computation as a matter of operations on constituent structure must be supplemented by an account of the local processing (which I'll loosely call computation) that perceptual icons undergo. A prominent iconophobe puts the problem succinctly: "This is a strong conclusion about cognitive architecture. It says, in effect, that the symbolic code idea that forms the foundation of computational theories does not apply to mental images" (Pylyshyn 2002, 162). This is the problem of local processing.

1.2—The binding problem(s). Our minds are more than a collection of atomic representations of distinct objects, properties, and relations. A typical human being can think not only about squares and about the color red, but also about red squares. The content <that is a red square> is somehow made up of the contents <red> and <square>. There must be some relation between the representations that express these simple contents that enables representation of the complex content. Explaining how contents bind together into a complex but unified whole is sometimes known, especially in perceptual psychology and neuroscience, as the binding problem (Clark 2004). It is also known as the feature integration problem (Treisman and Gelade 1980) and the many properties problem (Jackson 1977).

The binding problem comes in various forms. One is the *neural* binding problem, which concerns how representations in disparate neural areas come together (Feldman 2013). Another is

the *phenomenal* binding problem, the problem of accounting for how conscious experience can be made up of separate elements but nonetheless feel phenomenally unified (Bayne 2010). A third is the *computational* binding problem, the problem of understanding which lower-level computational algorithms are implemented by subpersonal mental systems in order to bind separate person-level vehicles or contents together.

The two forms of the binding problem that will be at issue here are the *syntactic* and *semantic* binding problems. The syntactic binding problem is the problem of how a single representational vehicle can have multiple contents. The semantic binding problem is the problem of how a single content can incorporate multiple contents together.

1.2.1—The semantic binding problem. The semantic binding problem is closely related to a problem philosophers of language call the problem of the unity of the proposition (Gaskin 2008; Jespersen 2012). I don't wish to engage directly with that problem here. For now, we can construe the semantic binding problem innocently enough that *predication* provides an answer to it. That is, the content <that is a red square> binds <red> and <square> together because both simple contents are predicated of the referent of <that>. This fact solves the semantic binding problem, but may not be anywhere near sufficient to account for the unity of the proposition—perhaps that problem may be restated as a problem of what semantic predication[45] really amounts to.

In any case, the semantic binding problem arguably arises for icons. It is not clear that icons are capable of expressing predication (see the following subsection for explication of the notion of semantic expression). Predication may require a logical apparatus that icons lack. In particular, it

---

[45] The phrase 'semantic predication' may seem prolix, but I will shortly contrast semantic with syntactic predication.

may require a separation between arguments and predicates that does not figure in the proper semantic analysis of icons. But icons surely bind contents together: a mental image of a red square combines <red> and <square> such that an image of a red square and green triangle differ semantically from an image of a red triangle and a green square despite the fact that those two images share the same simple contents <red>, <green>, <square>, and <triangle>.

The extent to which one accepts the existence of a semantic binding problem for icons will likely co-vary with the extent to which one accepts the conceptual coherence of so-called "nonconceptual content" as semantically different from the "conceptual content" of propositional thoughts. One might think that all content is propositional, and think it "decidedly puzzling" (Byrne 2005, 245) to suppose that perception could relate us to propositions that we could not ever be able to think (i.e., that are nonconceptual). On Russellian views of propositions, or on Stalnakerian possible-worlds views of content (Stalnaker 1984), it is indeed hard to see how any content graspable by perceptual systems could be impossible to express by means of words or concepts.

Philosophers of perception interested in nonconceptual content (e.g., Evans 1982; Peacocke 1992; Heck 2000; Burge 2010) tend to endorse Fregean views of propositional contents, on which the contents of thoughts are Fregean senses and the contents of perceptual states might be something else entirely (such as Russellian or Stalnakerian contents, or even Fregean senses of a distinctly nonpropositional sort—see Burge 2010). On such Fregean views, semantic predication can be understood as a Fregean sense or perhaps as a relation between Fregean senses, and its realization in a representation may depend upon certain structured conceptual or linguistic

capacities. In that case, a lack of structure may prevent perceptual icons from facilitating the grasping of predicative Fregean propositions. And in *that* case, icons cannot solve the semantic binding problem by means of semantic predication. We may need some alternate account of the semantics of icons that enables them to bind multiple contents together without semantic predication.

1.2.2—The syntactic binding problem. One may reject the Fregean framework required to formulate the semantic binding problem. The syntactic binding problem, however, cannot be sidestepped so easily. Consider the thought THAT IS A RED SQUARE. This thought is a complex mental representation partially composed of the concepts RED and SQUARE. These vehicles are syntactically bound together into a larger structure in a way that enables their contents to be predicated of the referent of THAT. This syntactic concatenation mirrors the semantic predication found in the propositional content of the thought; I will thus refer to it as *syntactic* predication. Syntactic predication is arguably what marks the format of thought as propositional (Camp 2007, 157).

Whether or not the best way of understanding the contents of icons is in terms of semantically predicative propositions, icons do not achieve syntactic predication. In an iconic representation of a red square, the part that represents its red color, the part that represents its square shape, and the part that represents the individual that has those features are all *the very same part of the icon*. In the linguistic expression 'That is a red square', we can literally point to the part that represents the individual and to the parts that correspond to each predicate; in a photograph of a red square, on the other hand, there is only one relevant part of the picture to point at. The mental analogues of sentences and pictures—propositional thoughts and perceptual icons,

respectively—may not have parts one can point at, but they do have functionally individuated syntactic parts that generate the same sort of distinction. The concepts RED and SQUARE have separate representational vehicles that both inhere in the thought THAT IS A RED SQUARE, while an iconic mental image of a red square lacks a canonical decomposition into separate syntactic predicates, and thus expresses the contents <red> and <square> via one and the same part of the mental image. Furthermore, icons lack explicit vehicles that correspond to individuals that have represented properties—they do not come labeled with names or demonstratives.

Here is another way of putting the syntactic binding problem: What syntactic features do a representation of a red square and a green triangle and a representation of a green square and a red triangle have in common? And what syntactic features do they not have in common? When it comes to representations with constituent structure like propositional thoughts, the form the answer should take is clear. The conjunctive thought THIS IS A RED SQUARE AND THAT IS A GREEN TRIANGLE shares the concepts RED, GREEN, SQUARE, and TRIANGLE with the conjunctive thought THIS IS A RED TRIANGLE AND THAT IS A GREEN SQUARE. Those thoughts differ in how they connect those categorial concepts together as well as which demonstrative concepts they are connected to via syntactic predication. In an icon, however, the form the answer should take is not so clear. We cannot isolate separate compositionally efficacious syntactic constituents shared by an image of a red square and green triangle and an image of a red triangle and green square. Nor can predicate-like elements of an icon be connected to representations of individuals via syntactic predication. This is the syntactic binding problem for icons.

1.3—The problem of explicit content. Here's a picture of Barack Obama:

185

Figure 16—Obama

What is the representational content of this picture? It certainly seems to represent colors, spatial relations, luminance edges, and other low-level features. Does it represent Obama? Does it represent Obama as smiling? Does it represent his smile as emanating from a misguided optimism about his presidency? Does it represent the flags behind him as being occluded by his body? Does it represent them as being occluded by Michelle Obama's husband's body? Does it represent him as standing (even if he is, in fact, crouching)? Suppose the picture was not taken of Obama but was instead a photograph of a complicated art installation of pieces of colored metal oriented in such a way as to exploit depth cues and look like Obama from precisely this vantage point. Would it still represent the scene as containing Obama?

There is a difference between what a representation explicitly represents and what it represents only implicitly. The sentence 'Barack Obama was born in Hawaii' represents the 44th president of the United States, but not explicitly so; it doesn't represent him *as* the 44th president. It does, however, explicitly represent him as Barack Obama as well as being born in Hawaii.

Discursive representations, particularly ones that closely resemble language, provide a somewhat straightforward way of delineating the explicit content of a complex representation. The predicate-constituents and their argument-constituents in a sentence or propositional thought have

186

semantic values; the explicit content of the whole representation is a function of the semantic values of each constituent together with the way the constituents are syntactically connected in the representation. Knowing what 'Barack Obama', 'was', 'born', 'in' and 'Hawaii' mean, together with how they are connected in the sentence, enables us to know straight away what the explicit content of 'Barack Obama was born in Hawaii' is. One needs the additional piece of information that Barack Obama is the 44th president of the United States to know that the sentence represents the 44th president as being born in Hawaii. Explicit content can thus be delineated by appeal to the semantic values of constituents and the semantically relevant relations between those constituents, while implicit content must be inferred from other information.

In the case of icons, it seems reasonable to search for a distinction between explicit, "bare-bones" (Haugeland 1998, 185) content and implicit content, including, arguably, high-level content (cf. Kulvicki 2006, Chapter 6). But the story just told about discursive representations cannot apply to icons. Icons, to repeat, lack constituents. It follows that the explicit content of an icon cannot be a function of its constituents and their relations to each other.

One might think this does not present a special problem, since icons are not the only symbols that lack constituents. Syntactically atomic symbols can have propositional contents—for example, a 'T' on a zoo enclosure might have the content <this is a tiger enclosure>. So perhaps there is a more general way of explicating content without constituents that abstracts away from iconic format. The prospects of this general approach are grim, however. Since atomic symbols lack *any* sort of combinatorial syntactic apparatus, each atomic symbol requires its own independent semantic interpretation. To apply this model to icons, each icon would have to acquire its content

independently—slightly change the color of one tiny part of an icon, and you have to have a completely new semantic interpretation for the whole icon. This approach is implausible on its face, and it leads to an explosion of primitive semantic values. I'll try to show later that it can be avoided, but it should in any case be clear that we should try to have an account of the syntax-semantics interface of icons that enables some sort of semantic compositionality.

We might just throw our hands up and say there is no way of delineating explicit content in an icon, and that Figure 1 represents Michelle Obama's husband's optimism just as well as it represents some pattern of light and dark. Fodor (1981a; 2007) seems to tend in this direction. Following Wittgenstein's puzzlement that a picture of a man walking uphill could just as easily depict a man walking downhill backwards, both Fodor (1981a) and Kosslyn (1980) argue that both pictures and perceptual icons lack contents except via associated propositional descriptions. This kind of surrender seems like a severe overreaction, especially given that examples like Wittgenstein's are plausibly about implicit high-level content and do not undermine the idea of explicit "bare bones" content. We might instead say that the explicit content of an icon is a matter of artistic intentions or conventions of interpretation. But this kind of approach, while it may be correct for at least some semantic aspects of artifactual icons, will not apply to mental representations like perceptual icons. We therefore need a way of delineating the explicit contents of icons that does not face either of these pitfalls. This is the problem of explicit content.

This problem is not of merely theoretical significance. We live in an increasingly iconic world. The ease of creating, reproducing, and transmitting images means that a larger portion of the representations we use to communicate are iconic than they have ever been in recorded history.

The question of how icons should be constructed and interpreted is of communicative, moral, and even legal significance.

Consider the case of emojis: small, relatively simple icons used to convey meanings in text messaging, email, social media, and other electronic forms of communication. The Washington Post reported in early 2016 that "a growing number" of cases have arisen "where authorities contend the cartoonish symbols have been used to stalk, harass, threaten or defame people," which has "left the police and courts wrestling with" how to interpret the content of an emoji.[46] For example, the Post mentioned that a New York City grand jury was forced to determine whether a use of the emojis in Figure 17 (in the depicted order) constituted a threat to the safety of police officers. While the case against the 17-year-old who used these emojis on social media was thrown out by the grand jury, the president of the Patrolmen's Benevolent Association wrote that the teenager "sent a message through social media that was easily interpreted by any reasonable person as a call for violence against police officers."[47]

Figure 17—A threat?

No doubt these cases must appeal to the general conventions surrounding the use of emojis and, especially, the recoverable intentions (especially communicative intentions) behind particular

---

[46]   http://www.washingtonpost.com/news/local/wp/2016/02/27/a-12-year-old-girl-is-facing-criminal-charges-for-using-emoji-shes-not-alone/

[47]   http://www.nydailynews.com/new-york/nyc-crime/grand-jury-tosses-brooklyn-teen-emoji-threat-charges-article-1.2101735

uses. And the most important practical, moral, and legal issues might not be the explicit content of an emoji but rather the kind of speech act it may be used to perform. Nonetheless, these factors are doubtless shaped by the explicit content of an emoji. In Figure 2, for example, it seems to matter that the gun emoji is "pointing" in the direction of the police officer emoji in determining whether its use expressed an intention to convey bodily harm to police officers. In this sort of case, specifying the low-level, explicit content of the icon is important for practical as well as theoretical purposes.

One might think these issues are straightforward and that we can easily translate from propositional to iconic forms of representation and back again by relying on general semantic competence. But Miller et al. (2016) found that, while each emoji has a corresponding Unicode description such as "grinning face with smiling eyes," these underlying descriptions are rendered into subtly (and sometimes drastically) different icons across different platforms in ways that lead to strikingly different interpretations from users. For example, Figure 18 shows the Google and Apple emojis corresponding to the Unicode description 'grinning face with smiling eyes'; the Google icon was rated much more positively than the Apple icon, and the former was described by a participant as "blissfully happy" while the latter was described by a participant as "ready to fight." While there perhaps may be some semantic incompetence at Apple, it seems that specifying the explicit content of an icon is not so easy.

Figure 18—"Grinning face[s] with smiling eyes" as interpreted by Google (left)

and Apple (right). Miller et al. (2016) found that the Google image was rated much more positively.

I focus on emojis only because they provide an emerging case where the proper understanding of the semantics of icons is very much up for grabs, and in a way that is of some social importance. As our representational media become less explicitly linguistic and more iconic, we should turn philosophical attention to the meanings of the symbols we use.

1.4—Signposting. These problems—the problem of local processing, the syntactic and semantic binding problems, and the problem of explicit content—are solved for discursive representations by appeal to their constituents. One and the same syntactic feature of discursive representations enables them to figure in local processing, to bind contents and vehicles together, and to distinguish explicit from implicit content. This kind of unifying explanation is elegant and satisfying. I will likewise appeal to a feature of the syntax of icons to solve these problems in a similar way, but without constituents. Icons, I'll argue, are syntactically and semantically analog: each part of an icon simultaneously holds values along multiple continuous syntactic dimensions that correspond to continuous dimensions of represented features, and thereby solve the problems mentioned without having constituent structure.

§2. Feature Placing

One might look at the suite of problems just articulated and conclude that there must be something wrong with the notion of iconic representation. We might instead search for an alternative syntax and semantics of representations in mental imagery and early perception (hereafter "sensory representations"). Clark (2000, 2004) pursues just such a strategy. Clark describes sensory

representations in terms of a *feature-placing* model. Feature-placing models (also endorsed, along roughly the same lines, in Tye 1991) construe sensory representation as a matter of "placing" features "at" locations.

The metaphor suggests both a semantic and a syntactic reading. On the semantic reading, the idea is simply that the content of a sensory representation is that an instance of some feature is present at a location or placetime. "Placing" is understood in terms of attribution; for a feature to be placed at some location is for it to be attributed to that location. On the syntactic reading, feature-placing models quantify over an array of mental "locations" at which symbols are "placed"; symbols stand for different features, and a feature is predicated of a location if a symbol corresponding to that feature is placed at the mental "location" corresponding to that objective location (Tye 1991).

It is not always clear which reading Clark intends. He is certainly committed to the truth of the semantic reading (e.g., 2004, 450), but he also posits mental sensory qualities corresponding to the objective features attributed to objects (a view developed in Clark 1993). While this latter point seems to suggest a syntactic reading, Clark (1996) adamantly rejects the idea that there are mental analogues of location which could be syntactically concatenated with mental sensory qualities. He writes instead that his model commits to a "stand-in for the capacities of spatio-temporal discrimination in virtue of which the sensory quality appears where it appears" (2004, 453). I confess I find this notion difficult to grasp without reading it as an appeal to mental location. Perhaps Clark does not mean to specify the syntactic structure of sensory representation at all, in which case his account may be compatible with propositional as well as non-propositional formats.

Indeed, his apparent agreement with Pylyshyn with respect to the imagery debate (Clark 2009), differing only on whether visually represented individuals are exclusively objects or also include locations, suggests that his feature-placing model is friendly to propositional format. Tye's (1991) discussion of symbol-filled arrays seems more clearly to commit to mental locations syntactically connected to mental representations of features.

I will suppose that, minimally, a syntactic reading of feature-placing requires some such relation between separate representations for locations and for features. I will also put aside purely semantic feature-placing accounts (though I will return to the semantics of feature-placing later) and will hereafter take feature-placing to be a syntactic proposal.

Feature-placing models are importantly different from iconic models of sensory representation (Clark 2009; Prinz 2011). Iconic models are marked by the lack of canonical decompositions in the syntactic structures of the representations they posit. Feature-placing models, on the other hand, are awash in canonical decompositions. A particular instance of feature-placing will decompose into a mental representation of a feature, a mental representation of a location, and a syntactic relation between them that makes the larger structure semantically evaluable in terms of the represented feature being instantiated at the represented location.

One potentially worrisome aspect of this feature-placing model is that it seems to implicate a propositional format and cannot therefore specify the respects in which sensory representations differ from thoughts. First of all, it requires a syntactic canonical decomposition. But more than that, the syntactic canonical decomposition mirrors the semantic canonical decomposition of a

proposition.[48] In this respect, the format of a feature-placing complex seems to be importantly propositional. The representation decomposes into a constituent corresponding to a predicate (i.e., a feature) and one corresponding to an argument (i.e., a location), and the syntactic relation between these functions to express the semantic relation of predication. Exhibiting this kind of syntactic predication is arguably sufficient (perhaps even necessary) for a representation to be syntactically propositional (Camp 2007, 157).

The apparently propositional representations posited by feature-placing models may nevertheless differ from those in central cognition in other sorts of ways. For example, relations between mental representations of location in sensory systems may mirror relations between actual or imagined locations (Tye 1991), while cognitive representations of location may differ from each other arbitrarily (just as linguistic representations of locations typically do). And mental representations of features in sensory systems may stand in quality-space relations to each other (Clark 1993; Rosenthal 2005; 2010), while cognitive representations of features need not. The worry about feature-placing models adverting to propositional constituent structure may therefore be consistent with representations in feature-placing systems differing from propositional thoughts in other syntactic ways.

Nonetheless, I think the appeal to proposition-like structure that characterizes feature-placing models is unmotivated. The main problem Clark appeals to his feature-placing model to

---

[48] This point obviously assumes a structured view of propositions; unstructured propositions won't have canonical decompositions. Nonetheless, one can still distinguish propositional format on an unstructured view of propositional content by appeal to the fact that a propositional representation exhibits a predicate-argument structure while rejecting the claim that this structure maps onto any semantic structure in the proposition(s) expressed.

solve is the binding problem (e.g., Clark 2004). In short, features can be bound by attributing them to placetimes—and mental representations of features can be bound by syntactically "placing" them at placetimes (though, again, I'm not confident that Clark should be interpreted this way). I will now put forward a rival model: the *coordination model* of sensory representation. The coordination model has much in common with Clark's feature-placing model, in that it appeals to nonconceptual states that have mental quality spaces. Unlike the feature-placing model, it appeals to iconic representations and lacks a predicative syntax and, depending on how the chips fall on more general semantic issues, it may even lack a predicative semantics.

Perhaps some feature-placing model like Clark's is in fact better suited to explain sensory representation on empirical grounds. Nonetheless, since such models do not involve iconic representation, we still need an account of how icons work. If Clark's model is better suited to explain human perception, it is not because icons are *conceptually* incapable of expressing contents, binding features, and figuring in local computational processing. The coordination model is thus worth developing whether or not it accurately describes human sensory representation. Furthermore, since the main motivation for Clark's model is its solution of the binding problem(s), showing how icons can do so without feature-placing goes some way to undermining the motivation for feature-placing models.

I will first sketch the coordination model before showing how it solves the problems for icons highlighted above.

§3. Analog Values and the Coordination Model

3.1—Analog syntax and semantics. There is an intuitive, but theoretically vexed, distinction between *analog* or *continuous* representations and processes on the one hand, and *digital* or *discrete* representations and processes on the other. The intuitive extreme cases of analog and digital representation are pictures and words, respectively. On the processing side, the distinction between analog and digital computation is well-known and arguably quite old (MacLennan 2012). In cognitive science, classical computational theories of mind posit discrete representations and computational rules sensitive to their structural features while, as noted above, operations in mental imagery such as mental rotation and image scanning have suggested a need for alternative, perhaps analog, models (Shepard and Cooper 1982).

Nelson Goodman (1976, Chapter 4) argued that the key to analog systems of representation is syntactic and semantic "density." A system is syntactically dense if there is always a third character (or symbol) between any two characters that differ syntactically, and a system is semantically dense if there is always a third character between any two that differ semantically. Goodman's requirement that analog symbol systems be both syntactically and semantically dense is too strong, and misses out on real joints between analog and digital representational systems. According to Goodman's criterion, a mercury thermometer is genuinely analog only if every possible change in the degree of height makes both a syntactic and semantic difference. A thermometer in which small degrees of height are irrelevant, and only hitting certain thresholds is syntactically and semantically significant, counts as "notational" (i.e., digital and language-like) by Goodman's criterion. We should thus be wary of the idea that analog representations must vary in

perfectly continuous ways. To borrow a similar example from Pylyshyn (1984, 200), the second hand on an analog watch intuitively represents time in an analog fashion whether it glides smoothly or tics from second to second.

I think the philosophical literature trying to spell out *the* uniquely correct criteria for distinguishing analog systems (e.g., Lewis 1971; Kulvicki 2015a) is a bit misguided. The word 'analog' could be used to refer to a type of computation (e.g., Anonymous 1962/2008; MacLennan 2012), to what I call iconic format (e.g., Carey 2009; Shea 2011; Kulvicki 2015a), to the use of magnitudes to represent magnitudes (Lewis 1971), to the representation of non-nested contents (Dretske 1981), or to a number of other interesting features of representations. All that matters is that a particular author's usage is clear, consistent, theoretically useful, and refers to something real; beyond that, different usages may cut representational systems at different, equally legitimate joints.

As I will use the term, iconic representations are syntactically analog in the sense that parts of icons take values along various continuous (though not necessarily dense) syntactic dimensions. This analog aspect of the syntax of icons enables[49] icons to represent such that (i) altering the iconic part along a certain syntactic dimension will alter the content of the icon along a corresponding semantic dimension, and (ii) the degree of semantic alteration will shift to a degree that corresponds in a lawlike way to the degree of syntactic variation.

---

[49] I write 'enables' rather than 'requires' because an artist may choose not to exploit the analog syntax of icons in representing what she aims to represent. The factors that determine the contents of artworks, such as the mental states of artists and interpreters, may be quite independent of the kind of bare syntactic and semantic apparatuses I articulate here (e.g., Briscoe 2016). The nature of photographs—specifically, that artistic intentions do not mediate their contents as directly or completely as other iconic forms of art—may require that their bare-bones content satisfies the analog semantic conditions laid out here. A painter, on the other hand, might decide to represent a scene in a way that makes very little use of the analog syntax of icon. If this line of thinking is right, then the analog syntax of icons makes available, but does not force, an analog bare-bones content. I will discuss this more in the conclusion.

Take any part of a photograph. That part will have a value along multiple color dimensions such as hue, saturation, and brightness.[50] These values are syntactic, since they are semantically exploitable, locally individuated features of the representational vehicle. Furthermore, they vary in a continuous fashion: a part of a photograph could be a little brighter or a lot brighter, or somewhere in between. Furthermore, these syntactic dimensions correspond to semantic dimensions; the brightness value in a part of a photograph corresponds to a brightness value in what it represents. In the case of photographs, the syntactic dimension can be the same as the semantic dimension to which it corresponds—viz., in the present example, brightness.[51] But that is not a necessary feature of analog representation. A mercury thermometer has a syntactic dimension of height along which it varies, and its syntactic value on this dimension corresponds to its semantic value on the dimension of ambient temperature. A functional icon, such as a mental image, can have functionally specified syntactic dimensions that map onto the discriminatory capacities that the representation facilitates, such as functional spatial dimensions (Kosslyn 1980, 1994).

It does not matter for present purposes whether the syntactic or semantic dimensions of an icon are dense. Consider again Pylyshyn's example of a second hand that tics as opposed to gliding from second to second. Though each tick is discrete, and there may be no syntactic or semantic continuity between two adjacent ticks, still the movement of the second hand is nonetheless a matter of degree—it can move more or less, and some locations are closer to each other than others.

---

[50] I intend to sidestep controversy about the precise dimensions one should use in characterizing color space. However that shakes out, it will not affect the point at issue. I will refer to hue, saturation, and brightness, but color aficionados may mentally substitute their preferred dimensions, such as lightness or chroma, at will.

[51] This sort of case, where a representation exemplifies the property it represents (and represents that property in virtue of exemplifying it), is what Peirce (1894) originally had in mind in coining the term 'iconicity'.

Furthermore, the degree of distance between two locations of the second hand corresponds to the degree of distance between the times represented at those locations. A second hand that moves from the two to the three and then to the four is getting syntactically and semantically further from its original location.

This is the sense in which these representational schemes are analog: they vary syntactically by degrees in such a way that values may be more or less close to each other, and the distance between these values tracks distances between the properties represented. Contrast the partially discrete second hand with a digital system that simply writes down numerals between 1 and 60, every second erasing the previous numeral and writing the succeeding numeral in its place. In this case, the move from numeral 2 to 3 to 4 is representing increasing temporal values, but there is no syntactic distance relation between the numerals that tracks represented temporal distance. This numeral system is in no sense analog.[52]

Kulvicki (2015a) describes a case of a mercury thermometer whose marks are arbitrarily assigned temperature values such that there is no interesting relation between how high the mercury is and how high the temperature. This odd thermometer is syntactically analog, since the level of mercury varies in degrees, but is not semantically analog, since differences in temperatures represented at different marks will not track the degree of syntactic variation (i.e., the degree of height in the mercury). A representational system can thus be syntactically analog if its vehicles

---

[52] This claim is, strictly speaking, false. The token numerals in the case described would stand to each other in temporal relations that would suffice for some form of syntactic analogicity. More precisely, the digital clock is analog in respect of its syntactic temporal properties but not its syntactic spatial properties, whereas the analog clock is analog in respect of both its spatial and temporal properties. I gloss over this point since any system that updates its syntactic representation of passing seconds each passing second would be analog in respect of its syntactic temporal properties. This rather deep point about continually representing time as it passes should not affect the larger point at issue.

vary in a continuous fashion, but it may fail to have an analog semantics if distance relations between syntactic values do not track distance relations in what is represented. Representational systems that are syntactically but not semantically analog are typically redundant, since they fail to exploit their own expressive power and could instead be more efficiently replaced with discrete representations—Figure 19 provides a helpful (and funny) example of the fecklessness of such systems.
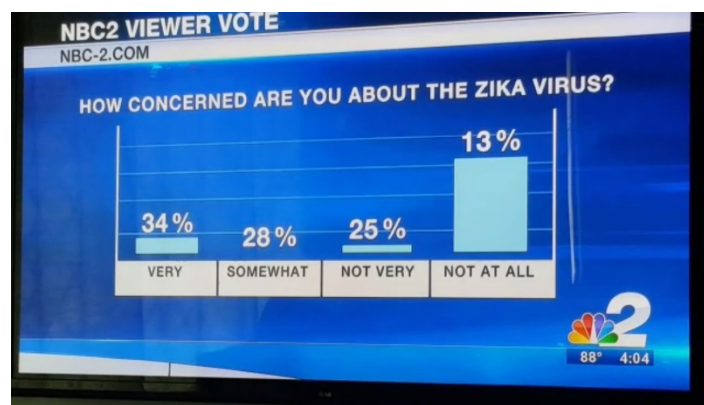
Figure 19—An inept graph used on the news (taken from reddit.com). Note that the failure to semantically exploit the syntactic analogicity of the bar graph renders it useless.[53]

The focus in the literature on the idea of perfect continuity or "density" is misplaced. Analog values can be continuous in the sense just described even though the move from degree to degree is discrete in substantive ways. Goodman's density criterion misses out on the joint in nature between systems like the discretely ticking second hand and systems like the machine that writes and erases numerals. The former is analog, and indeed represents in a syntactically and semantically *continuous* fashion, without being syntactically or semantically dense.

---

[53]Taken from
https://i.reddituploads.com/09d4079fd0bf453586b8524478aac4fd?fit=max&h=1536&w=1536&s=0d63d22eed3d44a4 1002007990acdf2c, accessed on April 7, 2017.

The notion of distance between syntactic and semantic degrees in analog representation can be, but need not be, literal spatial distance. Color values can be closer or further from each other in respect of hue; points on an audio recording can be closer or further from each other in respect of time.

Two points will matter especially in what follows. First, even the spatiotemporal syntactic properties of icons are syntactically and semantically analog. A given part of a photograph will not only have color values, it will also have a value along the spatial x- and y-axes of the photograph. These values are just as analog as any others just discussed. One can move along a spatial axis a little bit, a lot, or in between, and points can be syntactically (indeed, literally) closer or further from each other. Furthermore, syntactic spatiotemporal differences correspond proportionally to spatiotemporal differences; a spot on the far right of a photograph represents a part of the scene that is further (along some spatial dimension) from a spot on the far left of the photograph than from a spot in the center of the photograph.[54]

Second, the same part of an icon can simultaneously have values along multiple syntactic dimensions. A single part of a photograph will have values along the hue, brightness, and saturation color dimensions—the *very same part* has these values *simultaneously*. Indeed, that part will also have values along spatial dimensions as well. This point will form the basis of the coordination model of iconic representation.

---

[54] There is a wrinkle here, in that photographs intuitively tend to represent three-dimensional scenes despite being syntactically two-dimensional. Thus two spots right next to each other on a photograph might represent parts of the scene that are extremely far apart on the z-axis of the scene (think of a photo of the Grand Canyon, where a part representing the ground right at one's feet might be right next to a part representing the bottom of the canyon).

3.2—The coordination model. Any part of an icon, as just noted, holds values along multiple syntactic dimensions simultaneously. The analog syntactic properties of a part of an icon can thus be specified by a set of coordinates, each member of which is a value along a particular syntactic dimension. These values, and relations between them, correspond to values along dimensions of represented properties. The content of a part of an icon, therefore, can be specified by a set of coordinates along dimensions of continuous represented properties, which will correspond to the set of coordinates along syntactic dimensions in that part of the icon.

For example, consider a very simple two-dimensional spatial icon with ten locations along the x-axis and ten locations along the y-axis. The syntax of a particular primitive part of the icon might be represented by the coordinates (4, 7, 89, 23, 70), where the first two coordinates are the spatial x- and y-axes, respectively, and the latter three are values along the dimensions of hue, saturation, and brightness, respectively.[55] Assuming a simple one-to-one mapping of syntax to semantics in this case (where the icon represents a two-dimensional scene with no relevant color constancy effects), this set of coordinates will accurately characterize both the syntactic and the semantic values of the relevant part of the icon. More formally, if a primitive part p of an icon has values along n syntactic dimensions, the compositional syntax of p can be given as an n-tuple $(i_m…j_o)$, where i_m denotes a value i along syntactic dimension m. The syntax of the icon as a whole will be accurately and exhaustively characterized by a (typically very large) set of sets of coordinates

---

[55] As has often been reiterated, every part of an icon represents some part of the represented scene. But as a matter of physical necessity, icons will have primitive parts. For present purposes, primitive parts are defined as the smallest parts of an icon that retain syntactic properties (which are to be understood in terms of coordination, as it is presently unpacked). One might think of primitive parts as pixel-like, though audio recordings and functional icons like sensory representations have primitive parts as well.

corresponding to its primitive parts. In a case where the icon does not exemplify the very properties it represents, there will be separate coordinates characterizing the syntactic and semantic values of a given part, and mapping functions from values along syntactic dimensions to values along semantic dimensions. There is a high degree of freedom in permissible kinds of mapping functions, and they might be one-to-many (as in the case of two-dimensional icons representing three-dimensional space, perhaps, or the case of shape emerging from color and locational properties).[56]

The coordination model specifies the syntax and semantics of icons in terms of sets of coordinates in multi-dimensional spaces incorporating continuous dimensions, including spatiotemporal as well as featural dimensions.[57] The guiding ideas of this account are by no means novel. As Haugeland writes, "iconic contents might be conceived as variations of values along certain dimensions with respect to locations in certain other dimensions" (1998, 192). The coordination model is influenced by the idea of quality spaces in Clark's (1993) work and Rosenthal's (2005; 2010) quality-space theory of the consciously experienced "qualitative character" of perceptual states. One of the most interesting and salient aspects of Rosenthal's quality-space theory is that, unlike Clark, he posits mental qualities for spatiotemporal properties as well as features like color, pitch, and volume. The same considerations that suggest that our sensory representations underwrite discriminative capacities by instantiating distance relations

---

[56] It also does not matter whether the spatial structure of an icon or its corresponding scene is Euclidean or non-Euclidean, or metric or topological, etc. All that matters, as has been emphasized, is that there are definable dimensions of variation.

[57] I use 'spatiotemporal' as a catch-all. It's not the case that an icon must represent both spatial and temporal properties. A photograph has spatial but arguably no temporal dimensions, an audio recording has temporal but arguably no spatial dimensions, and a video has both.

along featural dimensions apply to spatiotemporal properties just as well, so a quality-space treatment fits both kinds of properties equally (see Meehan 2002 for sustained discussion using Rosenthal's framework).

The coordination model incorporates key insights of Rosenthal's quality-space theory, but is in some respects quite different. Quality-space theory, like Clark's theory, is designed to account for aspects of conscious experience (though Rosenthal argues forcefully that mental qualities can occur unconsciously). In particular, Rosenthal argues that the conscious qualitative character of a given perceptual experience is determined by the quality locations determined by its modality-specific sensory elements when those elements are conscious (2005, Chapters 5, 6, and 7). Furthermore, according to theorists like Clark and Rosenthal, the continuous nature of perceptual representation is actually an account of the *psychosemantics* (Fodor 1987) of modality-specific perceptual representation, i.e., how it is that they represent things at all. This psychosemantic project has been taken up most explicitly by Berger (2015) for Rosenthal's quality-space theory.

The coordination model is not committed to either of these theses, and so does not entail the truth of either Clark's or Rosenthal's theory. Moreover, since the coordination model posits iconic representations, Clark's feature-placing model is incompatible with it; Rosenthal's quality-space theory is not obviously incompatible with it, but since he does not posit iconic representation, his theory does not entail the accuracy of the coordination model.

My aim thus far has just been to sketch the coordination model and its motivations without delving into details. I'll now turn to how the model deals with the problems for icons that led us here in the first place.

§4. De-problematizing Icons

Icons whose syntax and semantics are accurately described by the coordination model can solve the problems highlighted at the outset.

4.1—The problem of local processing. Recall that, unlike propositional representations, icons cannot be computationally manipulated in ways that are sensitive to their constituent structure, because they don't have any. This was the problem of local processing. But sensitivity to constituent structure, even on Fodor and Pylyshyn's (1988) classical computationalist theory, is not essential to local processing as such; rather, it is posited to explain systematicity, productivity, and signs of apparent structure sensitivity in inference (see Quilty-Dunn and Mandelbaum unpublished b).

On the coordination model, an icon has mathematically definable values along continuous syntactic dimensions. Systems can operate on icons by accessing these values (which are, *qua* syntactic properties, locally individuated and therefore locally accessible). Computational processes involving icons will operate according to rules that concern analog values along syntactic dimensions.

What rules the system follows in taking these values as inputs or delivering them as outputs, what information the system brings to bear in these computations, and how best to formally model the computational processes in question, will differ from process to process and system to system. My goal has not been to specify computational models of perception or mental imagery, but rather to sketch how it could be possible for transformations defined over representations that lack constituent structure to figure in computations. The syntactic apparatus of the coordination model enables us to make sense of the idea of "computation," i.e., local processing, for perceptual icons.

The computational upshot of the coordination model is also consistent with Shepard and Cooper's (1982) model of computational processes in mental imagery such as mental rotation. Shepard and Cooper argue that these processes exhibit a second-order isomorphism with respect to the physical changes they represent. The idea is that the linear relation between degree of rotation and time required to mentally rotate objects is due to the fact that stages of the computational process correspond in an isomorphic fashion to stages of actually rotated physical objects. Shepard (1978, 135) rightly points out that this sort of account does not *require* that the manipulated representations have an iconic format, or any other particular format. But representations with an iconic format as described by the coordination model would nonetheless *explain* the second-order-isomorphic character of imagistic computational processes. If the spatial properties of an imaged object are represented via analog values along syntactic functional-spatial dimensions as the coordination model specifies, then "rotating" the iconic mental image from one set of coordinates along those spatial dimensions to another will require running through the intermediate values along each dimension. This predicts and explains the fact that the computational process itself, which operates on these values, exhibits the second-order isomorphism described by Shepard and Cooper.

This result suggests that the coordination model gets something deeply right about the notion of mental imagery as it emerged in experimental psychology in the 1970s. The idea that manipulating a mental image involves continuous shifts of values along syntactic dimensions fits extremely well with the classic experimental literature on scanning and mental rotation. Furthermore, since the coordination model is intended to describe photographs and emojis as well

as mental images, its comfortable theoretical fit with mental imagery results suggests that theorists like Kosslyn were on the right track in characterizing mental images as sharing core syntactic features with literal iconic images. None of these factors provide direct deductive arguments in favor either of the coordination model of icons or of iconic models of mental imagery, but the surprising explanatory coherence and simplicity of these theories taken together provides powerful abductive reasons for thinking these approaches might come close to converging on the truth.[58]

4.2—The binding problem(s). The syntactic binding problem was essentially the problem of explaining syntactic composition without syntactic constituent structure. Or as it was also framed: What are the syntactic similarities and differences between an icon that represents a red square and blue triangle on the one hand and one that represents a red triangle and blue square on the other?

The answer provided by the coordination model is that primitive parts of icons bind features by holding values along multiple syntactic dimensions simultaneously. For photographs and other picture-like representations, shapes emerge out of complexes of primitive parts that coordinate colors and locations. Thus colors and shapes are bound together in a pictorial icon because the primitive parts of an icon coordinate syntactic analogues of colors and locations, and these primitive parts compose into a syntactic analogue of a colored shape. This aspect of the coordination-based solution to the syntactic binding problem mirrors the familiar Berkeleyan insistence that visual representation of shape is dependent upon visual representation of color (Berkeley 1710/2008, Part 1, Section 10, 86-87; see also Reid 1764/1997, Chapter 6, Section 8, 99).[59]

---

[58] All of this is *pace* Pylyshyn 1973; 1984; 2002; 2003; 2007.

[59] Note that agreeing with Berkeley about this does not require that color and shape are computationally derived together via the same subpersonal channels (which they in fact are not—see, e.g., Cavina-Pratesi et al. 2010). I take it

The semantic binding problem was the problem of how icons compose semantic values without semantic predication. Here the answer is more complicated. A possible answer, as aforementioned, is that cases of semantic predication outstrip cases of syntactic predication, in which case icons involve semantic predication just as well as propositional thoughts. Alternatively, however, the coordination model provides a notion of semantic coordination that does not necessarily involve semantic predication. I'll explore this idea in the discussion of nonconceptual content later on.

4.3—The problem of explicit content. The problem of explicit content was the problem of telling what an icon literally represents and what it represents only implicitly. The semantic apparatus of the coordination account provides a sketch of an answer: the explicit content of an icon is the set of sets of coordinates along the dimensions of low-level properties that map to the coordinates along syntactic dimensions instantiated in the icon. I call these properties "low-level" precisely because they need to vary along continuous dimensions, and while that condition applies to color, shape, location, pitch, volume, temporality, and other low-level properties, it arguably does not apply to high-level properties like being John Malkovich, being a pine tree, or even being a physical object (cf. Berger 2015). And indeed high-level perception as well as mid-level object perception do not appear to involve wholly iconic representations (Quilty-Dunn 2016).

What a representation explicitly represents is a reflection of its syntactic elements and their semantic values. In the case of discursive representations, these syntactic elements can be

---

that Berkeley's claim, like mine, is about the structure of person-level iconic representations (or "ideas"), not about the processes that generate them.

standalone constituents that pick out any individual, property, or relation. In the case of icons, the relevant syntactic elements cannot be constituents but must be analog values that map to values of continuously variable properties. An icon might also be given semantic interpretations that are not reflected in its syntax, but these will not be part of the *explicit* content of the icons. Icons are capable of explicitly representing only continuously variable properties.

§5. Upshots

The coordination model solves various heretofore outstanding problems with iconic format. It also has some upshots for independent issues about representation, perceptual and otherwise.

5.1—Systematicity. Fodor (1987) argues that (i) conceptual thought is systematic, and that (ii) this systematicity requires constituent structure. Conceptual thought is systematic in that being able to think JOHN LOVES MARY always goes along with the ability to think MARY LOVES JOHN. This ability is explained by the fact that there are separate syntactic items JOHN, MARY, and LOVES, which can be recombined according to built-in rules of syntactic composition. As Mohan Matthen notes, however, "if one is able to entertain, in visual imagination for instance, images of a blue circle and of a red square, then one is able to entertain visual images of a blue square and a red circle" (2005, 80).

The coordination model shows how a representational scheme that lacks constituent structure can nonetheless be systematic, albeit in circumscribed ways. If a part of an icon determines a certain set of coordinates along a given set of syntactic dimensions, it could in principle determine any other set of coordinates along those dimensions. In this sense, being able

to represent light red at location-1 and dark red at location-2 entails being able to represent dark red at location-1 and light red at location-2. Icons are systematic within the parameters delineated by their continuous syntactic dimensions.

This result undermines the immediate inference from systematicity of *content* (which icons display) to systematicity of *vehicles* (which requires freely recombinable syntactic constituents). This may seem to diminish the force of systematicity arguments for constituent structure in thought. But the kind of systematicity displayed by icons is sharply circumscribed: it only allows for systematicity among continuously variable property dimensions. The classic examples of systematicity involving explicitly represented individuals, predicates, and relations are not subsumable under the coordination model, and may still provide reason to posit discursive mental representations.

5.2—Determinateness. It is often assumed without argument that icons must represent in a maximally determinate fashion. Berkeley claimed to be unable to represent any object without determinately representing its low-level properties (1710/2008, Introduction, Section 10, 72). More recently, Dennett (1981) has argued against iconic models of mental imagery on grounds that mental images can represent determinably.

One might rope the coordination model in with this tradition. After all, the idea of continuous analog dimensions seems to suggest that the semantic values of any part of an icon will be extremely fine-grained. The coordination model does indeed furnish the resources to account for the high level of determinacy that often obtains in icons. Nonetheless, the model does not entail that an icon must represent in a very determinate fashion at all. As argued above, the notion of

analog syntactic and semantic dimensions does not require determinacy (or "density" in Goodman's terminology). All that matters is that the degree of variance along some syntactic dimension can be greater or lesser, and that the degree of variance along some objective property dimension corresponds to the degree of change along the syntactic dimension of the representation. This scheme can be implemented in a representation with very coarse-grained, highly determinable values. For instance, a very low-resolution image can represent spatial properties in an equally analog fashion as a high-resolution image; the difference in that case is in the level of determinacy (i.e., the number of syntactic and semantic values expressible in the icon), not in the format of the representations.

Showing that a perceptual state represents non-maximally-determinate properties does not, therefore, suffice to show that the format of the state is non-iconic. One must show the property is not represented in a (perhaps low-resolution) analog fashion.

5.3—Nonconceptual content. Icons coordinate represented properties together. How should we understand this from a semantic perspective? Is the content of a primitive part of an icon a proposition? Does it feature semantic predication? These questions are directly relevant to the idea of so-called *nonconceptual content* in perception, which putatively differs from the contents of conceptual states like beliefs in purely semantic respects (e.g., Evans 1982; Peacocke 1992; Heck 2000).

Content on the coordination account is, in crucial respects, unstructured. The content of a primitive part of an icon is simply a set of values along continuous dimensions of iconically representible properties. While it might prove useful to formally model these sets as ordered n-

tuples with each position corresponding to a particular dimension, there is nothing intrinsic to the syntax or semantics of an icon that demands this sort of model. Instead, the content is simply the set of values, without any sort of internal structure. A Stalnakerian might argue that this makes iconic contents of a piece with propositional content more generally; content on such a view is an unstructured set of possible worlds (or functions from possible worlds to truth values). As such, perhaps the content of a primitive part of an icon should be understood as the set of possible worlds wherein the represented property values are co-instantiated. In that case, as Stalnaker (1998) noted in a different context, the distinction between conceptual and nonconceptual content is nil.

Let's suppose, however, that the contents of propositional attitudes are structured. What do we say about iconic contents given that presupposition? Perhaps that alone would justify distinguishing (structured) propositional contents from (unstructured) iconic contents. More than that, however, the lack of structure in iconic contents seems to preclude an analysis in terms of semantic predication.

In a discussion of contents as specified by his feature-placing model, Clark writes, "Specification of the content of an act of sense requires pairs of the form $([q_1...q_n], [r_1...r_m])$, and (I argue) the pairing principle is analogous to the tie between subjects and predicates" (2004, 453), where the first element of the pair specifies featural "qualities" while the second specifies placetimes. There are two crucial difference between this semantic account of sensory representation, which appeals to semantic predication, and the coordination model, which doesn't. First, values along quality dimensions are taken to cluster semantically and be separated as a group

from placetimes. Second, and more importantly, contents on the coordination model are semantically *symmetric* in a way that contents on Clark's feature-placing model are not.

Clark claims that qualities are predicated of, or attributed to, locations. On the coordination model, however, there is no more reason to say that a particular color instance is attributed to a particular location than to say that the location is attributed to the color instance. There is nothing in the syntactic or semantic apparatus of icons to settle the issue in one direction or the other. The lack of a way of distinguishing predicates from their arguments suggests that (given certain more general semantic presuppositions) iconic contents do not feature semantic predication. Coordinated iconic contents are, and semantically predicative contents are not, symmetric with respect to their semantic constituents. This fact might lead us to distinguish *semantic coordination* from semantic predication as a genuinely distinct semantic kind. This distinction would provide a novel approach to nonconceptual content as distinct from conceptual content in purely semantic respects.

Michael Rescorla (2009a; 2009b) argues that maps do not feature semantic predication. Iconic representation, though it is arguably present in maps, is a more general kind and does not constitutively involve the additional syntactic or semantic apparatuses made available in cartographic representation. Nonetheless, it would be surprising if bare icons such as photographs involved a *more* elaborate semantic apparatus than maps. In that case, an argument that maps lack semantic predication is directly relevant to the question of whether icons in general feature semantic predication.

According to Rescorla (2009b), once a marker is introduced in a map, the property denoted by the marker is represented as *not* being instantiated at locations in the terrain corresponding to locations on the map where the marker is absent. For example, if a map of a zoo has a 'B' for bathroom at location L1 and not location L2, Rescorla suggests that the map semantically expresses the absence of a bathroom at L2. Since semantic predication does not seem to work this way (e.g., if I say there's a bathroom on your right, I don't thereby say that there's no bathroom on your left), Rescorla argues that maps lack semantic predication.

Blumson (2012) and Bronner (2015) have argued that this feature of maps is really an aspect of the pragmatic, rather than semantic, features of maps. For example, Bronner provides the example of a map with the title "Some Cities I Have Visited," which represents certain U.S. cities. Intuitively, it is not the case that a failure to represent a city on that map represents that city as not having been visited, since the map is simply a map of *some* cities visited. The fact that the putatively semantic feature of maps identified by Rescorla can be canceled in this way suggests that it may instead be a pragmatic feature.

Rescorla argues that the absence intuition he identifies is not pragmatic. He imagines a mapmaker who, rather than simply failing to place a bathroom-denoting symbol at the part of the map corresponding to L2, instead randomly punches holes in the map including the part corresponding to L2 (2009b, 193). Rescorla argues that, if it turned out that there was a bathroom at L2, we would be much more upset with the mapmaker who left the part corresponding to L2 empty than with the mapmaker who punched out that part of the map. The Gricean solution, he claims, cannot explain this asymmetry. But it can. The hole-punching mapmaker was not

214

expressing his intention to convey the information that there are no bathrooms at L2, but the other mapmaker either was expressing his intention to convey that information, or failed to cancel the implicature that a reasonable consumer of maps would be expected to pick up on.

However, Rescorla (2009a) also provides a detailed compositional semantics of cognitive point-landmark maps on which, he argues, semantic predication simply does not enter the picture. He argues that cognitive maps as understood in robotics and cognitive science do not display hallmarks of logical form such as *modus ponens* inference, and develops a semantics that does not appear to invoke predication.[60]

Though I have not offered a compositional semantics for icons, the notion of semantic coordination seems, similarly, to characterize the semantic properties of icons without explicitly invoking predication (though see the conclusion for remarks about how a semantics of icons might proceed). Furthermore, as noted above, a compositional semantics of icons is unlikely to invoke a more elaborate semantic apparatus than a compositional semantics for maps. In that case, if Rescorla's semantics accurately describes the relation between maps and their truth values and genuinely does not involve semantic predication, then there may be good *prima facie* reason to think icons lack semantic predication as well.[61]

---

[60] Rescorla (2009a) argues the semantics of maps he provides does not involve explicit appeal to predication. But it is not clear from the semantics he provides that the contents of maps do not include propositions, and hence predication, involving the presence of properties represented by markers at their corresponding locations. The fact that maps do not involve a syntactically explicit predicative apparatus does not entail that they fail to accomplish semantic predication, and while Rescorla argues convincingly for the former thesis, it is not clear that the mere absence of an explicit appeal to predication in his semantic model suffices to rule out the (otherwise natural) idea that maps exhibit semantic predication.

[61] I argue below, however, that the semantics of artifactual icons may be shaped by the mental states of artists and conventions of interpretation, and this fact may incorporate non-explicit semantic predication into the contents of artifactual icons.

One might instead subscribe to a view of content on which semantic predication dissolves into a Russellian proposition, which is simply an ordered set of referents (construed broadly to include properties, relations, and functions as well as individuals), or one on which semantic predication is simply the determining of sets of possible worlds. Someone wedded to either of these views might be inclined to deflate the notions of *proposition* and *predication* such that *any* representation that is evaluable for accuracy with respect to some possible state of affairs thereby expresses a proposition and—perhaps for that reason alone—accomplishes semantic predication.[62] On this deflationary view, theorists like Peacocke (2001) and Burge (2010) who insist that perceptual content is non-propositional and non-predicative but nonetheless accuracy evaluable would be conceptually confused. In that case, the fact that the semantic properties of icons (or maps, for that matter) are different in important ways from the semantic properties of explicitly predicative propositional representations would not undermine the claim that both involve propositional contents and semantic predication.

The differences between coordination and predication, however, can simply be restated in terms of modes of presentation, without supposing that modes of presentation are semantic entities (see, e.g., Fodor 1998). Instead, modes of presentation might be defined as *whatever is responsible for differences in cognitive significance* (in Frege cases, canonically), and that definition might be satisfied by syntactic forms. For instance, the difference between a concept of Clark Kent and a

---

[62] I suspect much of the disagreement between so-called "content conceptualists" (e.g., Byrne 2005; Speaks 2005, and arguably Stalnaker 1998) and "content non-conceptualists" (e.g., Heck 2000; Peacocke 2001; Burge 2010) is due to the fact that the former do and the latter do not wish to deflate semantic notions in this sort of way. It is hard to see on a view such as Stalnaker's, for example, how an accuracy-evaluable representation could fail to express a proposition, or how semantic predication and semantic coordination could really pick out distinct sorts of contents.

concept of Superman might be purely a difference in the syntax of the mental representations, in which case they constitute different modes of presentation while being semantically equivalent. If this Fodorian view is the right approach to solving Frege puzzles while preserving a purely referential semantics, then the syntactic differences between icons and propositional representations expounded in this chapter suffice to demonstrate that perceptual icons involve distinctly iconic, analog modes of presentation even if they do not involve distinct contents. The notion of nonconceptual content in this case might be most usefully understood in terms of modes of presentation rather than in terms of semantic content *per se*.

§6. Conclusion: Toward a Semantics of Icons

Icons are completely unlike linguistic, conceptual, and other forms of discursive representation. Icons lack constituent structures, and every part of an icon represents some part of the represented scene in virtue of coordinating values along continuous syntactic dimensions corresponding to values along continuous semantic dimensions. This enables icons to figure in local processes, to compose syntactically and semantically, to distinguish explicit from implicit content, to exhibit circumscribed systematicity, to represent at varying levels of determinacy, and perhaps to exhibit nonconceptual contents.

A full account of the nature of iconic representation would include a semantics for icons, minimally consisting of rules for generating accuracy conditions. I have not provided such an account here. I'll conclude by pointing in some possible directions for developing a semantics of icons.

There are multiple possibilities in specifying the conditions under which an icon or its primitive parts are accurate. A simple idea, suggested above, is that an icon is accurate iff the properties it semantically coordinates are co-instantiated. This simple idea can provide a framework for developing a semantics of icons. For example, for an icon whose primitive parts vary along syntactic dimensions a, b, and c corresponding to property dimensions d, e, and f, a primitive part that coordinates syntactic values (a_1, b_1, c_1) will be accurate iff corresponding properties d_1, e_1, and f_1 are co-instantiated in the relevant scene.[63]

A central issue for such an account is determining the relevant scene (as opposed to some qualitatively identical scene somewhere else). This issue could be resolved by indexically anchoring spatiotemporal coordinates. For example, in the case of sensory representations, the relevant spatial framework may be egocentrically anchored (Peacocke 1992). In that case, given that values along spatiotemporal dimensions are among those coordinated with values along other property dimensions, the accuracy conditions of a perceptual icon that represents in an egocentric framework would specify the organism's actual environment and not some other, otherwise qualitatively similar environment.

Another resolution could come from incorporating causal relations into the accuracy conditions of the icon. That is, an icon is accurate with respect to scene s iff the properties it semantically coordinates are co-instantiated in s and its coordinating those properties is

---

[63] Co-instantiation here means not merely that the properties are instantiated (e.g., that some color and some location are both present in the scene), but that they are instantiated together (e.g., that the location has the property of instantiating that color, or, equivalently for our purposes, that the color has the property of being instantiated at that location).

appropriately caused by *s* (or by the co-instantiation of those properties in *s*). Some such condition might be necessary to account for cases of misrepresentation that contingently happen to match the scene (Soteriou 2000). Note that, while this semantics would build causal relations to individuals into the accuracy conditions of icons, it does not (and could not) build those causal relations into the explicit content of icons.[64] The lack of any syntactic property of icons corresponding to individuals or causal relations to them means that the icon does not explicitly represent individuals (beyond instances of coordinated properties) or causal relations to individuals. In that case, the iconic semantics at hand posits a division between what constitutively figures in the accuracy conditions of an icon and what the icon explicitly represents.

Philosophers like Hopkins (1998), Greenberg (2013), and Briscoe (2016) offer accounts of the semantics of pictorial representation by appeal to factors like resemblance (Hopkins 1998 and Briscoe 2016) or sets of three-dimensional worlds indexed to perspectives that can be projected from the two-dimensional image (Greenberg 2013). The coordination model is not at odds with any of these approaches. The coordination model aims to characterize the syntax of icons and some aspects of its relation to the semantics of icons; the latter could be understood in terms of resemblance, projection, or something else. For instance, perhaps the way in which coordination along two-dimensional spatial axes in a photograph enables coordination along three-dimensional spatial axes in represented scenes could be understood in terms of Greenberg's projection model. Perhaps the notion of semantic coordination, appealing as it does to correspondences of analog

---

[64] This has the consequence that an account of perceptual content like Searle's (1983), on which causal relations between the represented scene and the representation itself are explicitly represented, cannot be cashed out in purely iconic terms.

values along syntactic and semantic property dimensions, could be understood as instantiating a form of resemblance, and as explaining other forms of resemblance (such as similarity of the perceptual states caused by a pictorial representation on the one hand and by the scene it represents on the other).

As aforementioned, there are many other aspects of the meanings of icons such as paintings, drawings, etc., that are not captured by the coordination model. For one thing, a painting can arguably represent individuals and high-level properties that do not vary along continuous semantic dimensions (properties such as, e.g, the purported historical inevitability of westward territorial expansion by the United States in John Gast's famous pro-imperialist painting "American Progress"—see Fig. 20). A complete semantics of such representations will far outstrip the resources of the coordination model, and these high-level contents (which may arise in part due to artistic intentions alone) are arguably much more important for understanding art than the bare-bones resources of the coordination model. And different sorts of iconic or quasi-iconic representations, such as graphs, maps, diagrams, and models, likely require individual semantic analyses that may fail to generalize across representational kinds.

Figure 20—"American Progress" by John Gast, 1872

Moreover, while the coordination model takes it as a bedrock fact that icons are *syntactically* analog, they are not necessarily *semantically* analog. It need not be the case that an artist semantically exploits the analog syntactic apparatus of icons in an analog fashion. In the case of highly abstract art, there may be no semantic content at all, or it may be very indeterminate. Even iconic artworks with clear, specific contents can fail to coordinate semantic properties. Imagine an artistic convention wherein some specific light shade of blue represents sadness, the next lightest shade represents joy, and the next lightest shade represents fear. Suppose an artist knows this system and creates a painting with patches of these three shades next to each other, and intended to make her audience realize that each patch represented its respective emotion. Suppose also that this convention is well-known enough that audience members did generally realize this, and interpreted the painting accordingly. It would seem that, in such a case, the painting represents sadness, joy, and fear in a semantically discrete (i.e., non-analog) fashion. While the painting syntactically coordinates values along color dimensions and shape dimensions, it does not semantically coordinate values along represented dimensions; instead, it represents discrete categories in a way

221

more akin to language. This toy example helps illustrate that the syntactically analog nature of icons facilitates, but does not require, semantically analog content.

A surprising consequence of this fact is that the principle that has often anchored the notion of iconic representation—that parts of the vehicle represent parts of the scene—turns out not necessarily to be exhibited by icons. Icons are constitutively syntactically analog; in cases where this syntactic analogicity fails to be semantically exploited, parts of the icon will fail to represent parts of the represented scene. Nonetheless, icons *qua* syntactically analog representations are constitutively such that they can express analog contents (i.e., be semantically analog). It is thus a constitutive truth about icons that parts of an icon can be semantically exploited to represent parts of the scene, not that they actually do. And in the case of mental representation, where there is no equivalent of a painter to bestow arbitrary contents on an icon, perceptual icons may be semantically analog as a matter of necessity. An answer to this question depends on answers to deeper questions about how mental representations have contents at all.

THE PERCEPTION–COGNITION BORDER

Conclusion

The central aim of this dissertation has been to characterize the outputs of perception in a way that sheds light on larger questions about the architecture of the mind. The main positive thesis that has been advanced is perceptual pluralism, according to which perception delivers both unconceptualized iconic representations and conceptualized discursive representations (as well as, perhaps, other representational forms). A detailed account has been provided of the syntactic and semantic properties of iconic representation as well as the respects in which certain perceptual representations satisfy conditions on conceptuality.

What are the consequences of this characterization of the outputs of perception for mental architecture? In particular, how should we think about the border between perception and cognition?

Chapter 1 sketched an architectural strategy for distinguishing perception from cognition that is consistent with perceptual pluralism. According to this strategy, perceptual systems implement processes that are stimulus bound and informationally encapsulated from cognition. This strategy also holds that cognitive systems do not exhibit both of these properties (though some might exhibit one, e.g., certain forms of demonstrative thought might be stimulus bound but unencapsulated and central modules might be encapsulated but not stimulus bound).

A full defense of this architectural strategy would require a dissertation in itself. The primary defense given here is an indirect one, viz., that the truth of perceptual pluralism rules out

representational approaches to the perception–cognition border and that it is fully consistent with perceptual pluralism that perceptual processes are essentially stimulus dependent and encapsulated from cognition.

Chapter 2 also provided two defenses of encapsulation. First, it argued that the information used by the object perception system was encapsulated from and inaccessible to central cognition, as shown by the fact that what we cognitively think of as an object and its conditions for persistence come apart from how we visually individuate and track objects. Object perception furthermore has been shown to directly contradict subjects' judgments in at least one experiment (Mitroff et al. 2005).

Second, Chapter 2 provided some *a priori* reason to think that early perceptual processes are encapsulated. If early perceptual processes operate on and output perceptual icons, then there is reason to think they cannot operate on discursive representations stored in cognition. Either the perceptual systems would have to translate cognitive representations into iconic formats, or there would have to be rules that specify discursive states built into perceptual systems. While one or both of these disjuncts may be true, the difference in format nonetheless provides some barrier to direct violations of encapsulation from cognition.

The positive proposals made here also place some constraints on how the architectural strategy should be pursued.Chapter 1 provided reason to think that even widespread modulation of perceptual processing by cognitively driven attention is compatible with the encapsulation of perception from cognition. Moreover, that chapter sketched dimensions of encapsulation; what

matters for the architectural strategy on offer is that perception is encapsulated from cognition to some explanatorily significant degree, and these dimensions may help adjudicate this matter.

Perhaps the most significant constraint this dissertation places on the architectural strategy is the need for substantial diachronic cognitive penetration. If the arguments in Chapters 2 and 3 are right, then concepts like PIANO are deployed by perceptual systems. There is no independent reason to think that the concept PIANO is innately accessible to the visual system. Given the probable scarcity of pianos in the ecologies of our evolutionary ancestors, there is independent reason to doubt that the concept PIANO is innately accessible to the visual system or indeed any other human mental system. The fact that the visual system can deploy the concept PIANO should therefore make us think that learning processes can push concepts like PIANO into the visual system over time. How perceptual learning works and to what extent cognition is involved in perceptual learning are questions left open by this dissertation. But there must be some process by which learned concepts become deployable by the visual system.

To fully pursue the architectural strategy, one must grapple with the many putative counterexamples to the encapsulation of perception offered in the philosophical and experimental literature. Nonetheless, the arguments made here provide a framework for how to pursue this strategy for differentiating perception from cognition.

The distinction between iconic and discursive representational formats has played a central role in this framework. I suspect that future research will show this simple dichotomy to be too coarse grained. Perhaps a richer and more accurate characterization of the syntax and semantics of perceptual representation will provide enough subtle distinctions to do away with these broad

categories entirely. For now, however, the syntax and semantics of perceptual representation offered

here may point the way forward to a better, truer theory.

Bibliography

Anonymous. (1962). Why analog computation? *National Security Agency Technical Journal* 7(3) (declassified, 2008), 85–91.

Ariely, D. (2001). Seeing sets. *Psychological Science* 12(2), 157–162.

Audi, R. (2015). *Rational Belief: Structure, Grounds, and Intellectual Virtue*. Oxford: OUP.

Aydede, M. (1998). Fodor on concepts and Frege puzzles. *Pacific Philosophical Quarterly* 79(4), 289–294.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition* 10(8), 949–963.

Balcetis, E., & Dunning, D. (2010). Wishful seeing: More desired objects are seen as closer. *Psychological Science* 21(1), 147–152.

Bayne, T. (2009). Perception and the reach of phenomenal content. *The Philosophical Quarterly* 59(236), 385–404.

Bayne, T. (2010). *The Unity of Consciousness*, Oxford: OUP.

Becker, M.W., Pashler, H., & Anstis, S.M. (2000). The role of iconic memory in change-detection tasks. *Perception* 29, 273–286.

Berger, J. (2015). The sensory content of perceptual experience. *Pacific Philosophical Quarterly* 96(4), 446–468.

Berkeley, G. (1710). *A Treatise Concerning the Principles of Human Knowledge*. In D.M. Clarke (ed.), *George Berkeley, Philosophical Writings* (Cambridge: Cambridge University Press, 2008), 67–149.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review* 94(2), 115–147.

Bliss, J.C., Crane, H.D., Mansfield, P.K., & Townsend, J.T. (1966). Information available in brief tactile presentations. *Perception & Psychophysics* 1, 273–283.

Block, N. (1983). The photographic fallacy in the debate about mental imagery. *Nous* 17(4), 651–661.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences* 12, 567–575.

Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research* 89(3), 560–572.

Block, N. (unpublished). *A Joint in Nature between Perception and Cognition without Modularity of Mind*. Book manuscript.

Blumson, B. (2012). Mental maps. *Philosophy and Phenomenological Research* 85(2), 413–434.

Bradley, C., & Pearson, J. (2012). The sensory components of high-capacity iconic memory and visual working memory. *Frontiers in Psychology* 3(355), doi: 10.3389/fpsyg.2012.00355.

Braine, M.D.S., & O'Brien, D.P., eds. (1998). *Mental Logic*. Mahwah, NJ: Erlbaum.

Brandom, R. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Brewer, B. (2006). Perception and content. *The European Journal of Philosophy* 14, 165–81.

Briscoe, R. (2015). Cognitive penetration and the reach of phenomenal content. In A. Raftopoulos & J. Zeimbekis (eds.), *Cognitive Penetrability* (Oxford: OUP), 174–199.

Briscoe, R. (2016). Depiction, pictorial experience, and vision science. *Philosophical Topics* 44(2), 43–81.

Brockmole, J.R., Wang, R.F., & Irwin, D.E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance* 28(2), 315–334.

Brogaard, B. (2013). Do we perceive natural kind properties? *Philosophical Studies* 162, 35–42.

Brogaard, B., & Gatzia, D. (2016). Is color experience cognitively penetrable? *Topics in Cognitive Science* 9(1), 193–214.

Bronner, B. (2015). Maps and absent symbols. *Australasian Journal of Philosophy* 93(1), 43–59.

Bruce, V., & Young, A. W. (2000). *In the Eye of the Beholder: The Science of Face Perception*. Oxford: OUP.

Bruce, V., Green, P.R., Georgeson, M.A. (2014). *Visual Perception: Physiology, Psychology, and Ecology*, fourth edition. New York: Psychology Press.

Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy* 4(1), 73–122.

Burge, T. (2010a). *Origins of Objectivity*. Oxford: OUP.

Burge, T. (2010b). Steps toward origins of propositional thought. *Disputatio* 4(29), 39–67.

Burge, T. (2014a). Reply to Block: Adaptation and the upper border of perception. *Philosophy and Phenomenological Research* 89(3), 573–583.

Burge, T. (2014b). Reply to Rescorla and Peacocke: Perceptual content in light of perceptual constancies and biological constraints. *Philosophy and Phenomenological Research* 88(2), 485–501.

Butler, A., Oruc, I., Fox, C.J., & Barton, J.J.S. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Research* 1191, 116–126.

Butterfill, S.A., & Apperly, I.A. (2016). Is goal ascription possible in minimal mindreading? *Psychological Review* 123(2), 228–233.

Byrne, A. (2005). Perception and conceptual content. In E. Sosa and M. Steup (eds.), *Contemporary Debates in Epistemology* (London: Basil Blackwell), 231–250.

Camp, E. (2004). The generality constraint and categorial restrictions. *The Philosophical Quarterly* 54(215), 209–231.

Camp, E. (2007). Thinking with maps. *Philosophical Perspectives* 21, 145–182.

Camp, E. (2009). Putting thoughts to work: Concepts, systematicity, and stimulus-independence. *Philosophy and Phenomenological Research* 78(2), 275–311.

Campbell, J. (2002). *Reference and Consciousness*. Oxford: OUP.

Carey, S. (2009). *The Origin of Concepts*. Oxford: OUP.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: beyond object files and object tracking. *Cognition* 80, 179–213.

Carrasco, M., & Barbot, A. (2014). How attention affects spatial resolution. *Cold Spring Harbor Symposia on Quantitative Biology* 79, 149–160.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience* 7(3), 308–313.

Carruthers, P. (2009). Invertebrate concepts confront the generality constraint (and win). In R. Lurz (ed.), *The Philosophy of Animal Minds* (Cambridge: CUP), 89–107.

Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Churchland, P.M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy* 78, 67–90.

Clark, Andy. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36(3), 181–204.

Clark, Andy. (2016). Attention alters predictive processing. *Behavioral and Brain Sciences* 29, 23–24.

Clark, Austen. (1993). *Sensory Qualities*. Oxford: Clarendon Press.

Clark, Austen. (1996). Three varieties of visual field. *Philosophical Psychology* 9(4), 477–495.

Clark, Austen. (2000). *A Theory of Sentience*. Oxford: OUP.

Clark, Austen. (2004). Feature-placing and proto-objects. *Philosophical Psychology* 17(4), 443–469.

Clark, Austen. (2009). Location, location, location. In D. Dedrick and L. Trick (eds.), *Computation, Cognition, and Pylyshyn* (Cambridge, MA: MIT Press), 281–302.

Cohen, M., & Dennett, D.C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences* 15(8), 358–364.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics* 27(3),183–228.

Cosmides, L., Barrett, H.C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *PNAS* 107(2), 9007–9014.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24, 87–185.

Crane, T. (1992). The nonconceptual content of experience. In his *The Contents of Experience* (Cambridge: Cambridge University Press), 136–157.

Craver, C.F. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon.

Darwin, C.J., Turvey, M.T., & Crowder, R.G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology* 3, 255–267.

Davidson, D. (1967). Truth and meaning. *Synthese* 17, 304–23.

Dennett, D.C. (1978). A cure for the common code. In his *Brainstorms: Philosophical Essays on Mind and Psychology* (Bradford), 90–108.

Dennett, D.C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, D.C. (1991). Real patterns. *Journal of Philosophy* 88(1), 27–51.

Dennett, D.C. (2006). The selfish gene as a philosophical essay. In A. Grafen and M. Ridley (eds.), *Richard Dawkins: How a Scientist Changed the Way We Think* (Oxford: OUP), 101–115.

Deroy, O. (2013). Object-sensitivity versus cognitive penetrability of perception. *Philosophical Studies* 162(1), 87–107.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18, 193–222.

Devitt, M. (1990). Meanings just ain't in the head. In G. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam* (Cambridge: Cambridge University Press), 79–105.

Devitt, M. (2006). *Ignorance of Language*. Oxford: OUP.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Dretske, F. (2000). If you can't make one, you don't know how it works. In his *Perception, Knowledge, and Belief: Selected Essays* (Cambridge: Cambridge University Press), 208–226.

Dretske, F. (2004). Change blindness. *Philosophical Studies* (120), 1–18.

Dummett, M. (1973). *Frege: Philosophy of Language*. Second edition, 1981. Cambridge, MA: Harvard University Press.

Durgin, F.H., Baird, J.A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review* 16(5), 964–969.

Duysens, J., Orban, G.A., Cremieux, J., & Maes, H. (1985) Visual cortical correlates of visible persistence. *Vision Research* 25 (2), 171–178.

Egan, A. (2008). Seeing and believing: Perception, belief formation and the divided mind. *Philosophical Studies* 140(1), 47–63.

Elga, A., & Rayo, A. (unpublished). Fragmentation and information access.

Evans, G. (1982). *The Varieties of Reference*. Oxford: OUP.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology* 44, 33–66.

Feldman, J. (2013). The neural binding problem(s). *Cognitive Neurodynamics* 7(1), 1–11.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fields, C. (2011). Trajectory recognition as the basis for object individuation: A functional model of object file instantiation and object-token encoding. *Frontiers in Psychology* 2(49), DOI: 10.3389/fpsyg.2011/00049.

Finke, R. A. & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 8(2), 142–147.

Firestone, C. & Scholl, B.J. (2014). Enhanced visual awareness for morality and pajamas? Perception vs. memory in 'top-down' effects. *Cognition* 136, 409–416.

Firestone, C. & Scholl, B.J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences* 39, 1–72.

Flombaum, J.I., Scholl, B.J., Santos, L.R. (2008). Spatiotemporal priority as a fundamental principle of object persistence. In B. Hood and L. R. Santos (eds.), *The Origins of Object Knowledge* (Oxford: OUP), 135–164.

Fodor, J.A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J.A. (1981a). Imagistic representation. In N. Block (ed.), *Imagery* (Cambridge, MA: MIT Press), 135–149.

Fodor, J.A. (1981b). Introduction: Some notes on what linguistics is about. In N. Block (ed)., *Readings in the Philosophy of Psychology, Vol. 2* (Cambridge, MA: Harvard University Press).

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Fodor, J.A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon.

Fodor, J.A. (2004). Having concepts: A brief refutation of the twentieth century. *Mind and Language* 19(1), 29–47.

Fodor, J.A. (2007). The revenge of the given. In B. McLaughlin and J. Cohen (eds.), *Contemporary Debates in Philosophy of Mind* (Oxford: Blackwell), 105–116.

Fodor, J.A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: OUP.

Fodor, J.A., & Pylyshyn, Z. (1981). How direct is visual perception? Some reflections on Gibson's "ecological approach." *Cognition* 9, 139–196.

Fodor, J.A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 3–71.

Fougnie, D., & Alvarez, G.A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature–store model. *Journal of Vision* 11(12), 1–12.

Frisby, J.P., & Stone, J.V. (2010). *Seeing: The Computational Approach to Biological Vision*. Cambridge, MA: MIT Press.

Gantman, A.P., & Van Bavel, J.J. (2014). The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition* 132, 22–29.

Gaskin, R. (2008). *The Unity of the Proposition*. Oxford: OUP.

Gauker, C. (2011). *Words and Images: An Essay on the Origin of Ideas*. Oxford: OUP.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilbert, D. (1991). How mental systems believe. *American Psychologist* 46(2), 107–119.

Goodman, N. (1976). *Languages of Art: An Approach to a Theory of Symbols*. 2nd edition. Indianapolis: Hackett.

Gordon, R.D. (2014). Saccade latency reveals episodic representation of object color. *Attention, Perception, & Psychophysics* 76(6), 1765–1777.

Gordon, R.D., & Irwin, D.E. (1996). What's in an object file? Evidence from priming studies. *Perception and Psychophysics* 58(8), 1260–1277.

Gordon, R.D., & Irwin, D.E. (2000). The role of physical and conceptual properties in preserving object continuity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(1),136–150.

Gordon, R.D., & Vollmer, S.D. (2010). Episodic representation of diagnostic and non-diagnostic object color. *Visual Cognition* 18(5), 728–750.

Gordon R.D., Vollmer S.D., & Frankl M.L. (2008). Object continuity and the transsaccadic representation of form. *Perception and Psychophysics* 70, 667–679.

Green, E.J. (2015). A layered view of shape perception. *The British Journal for the Philosophy of Science*, doi: 10.1093/bjps/axv042

Green, E.J. (forthcoming). On the perception of structure. *Noûs*.

Green, E.J., & Quilty-Dunn, J. (unpublished). What is an object file?

Greenberg, G. (2013). Beyond resemblance. *The Philosophical Review* 122(2), 215–287.

Gross, S. (2017). Cognitive penetration and attention. *Frontiers in Psychology* 8(221), doi: 10.3389/fpsyg.2017.00221.

Grzeczkowski, L., Tartaglia, E.M. Mast, F.W., & Herzog, M.H. (2015). Linking perceptual learning with identical stimuli to imagery perceptual learning. *Journal of Vision* 15(13), 1–8.

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe and L. Robertson (Eds.), *From Perception to Consciousness: Searching with Anne Treisman* (Oxford: OUP), 339–349.

Haladjian, H., & Pylyshyn, Z. (2008). Object-specific preview benefit enhanced during explicit multiple object tracking. *Journal of Vision* 8(6), 497.

Haugeland, J. (1998). *Having Thought*. Cambridge, MA: Harvard University Press.

Heck, R. (2000). Nonconceptual content and the "space of reasons." *The Philosophical Review* 109(4), 483–523.

Helmholtz, H. von. (1867). *Treatise on physiological optics, volume 3*. J. P. C. Southall (tr.), 1962, New York: Dover.

Henderson, J.M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General* 123(4), 410–426.

Henderson, J.M., & Siefert, A.B.C. (2000). Types and tokens in transsaccadic object identification: Effects of spatial position and left-right orientation. *Michigan State University Eye Movement Laboratory Technical Report* 3, 1–12.

Hill, C. (2008). Review of *Things and Places: How the Mind Connects with the World. Notre Dame Philosophical Reviews*, published online 2008.07.10. URL: https://ndpr.nd.edu/news/23626-things-and-places-how-the-mind-connects-with-the-world/

Hohwy, J. (2013). *The Predictive Mind*. Oxford: OUP.

Hollingworth, A., & Rasmussen, I.P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance* 36(3), 543–564.

Hopkins, R. (1998). *Picture, Image and Experience*. Cambridge: Cambridge University Press.

Hopp, W. (2011). *Perception and Knowledge: A Phenomenological Account*. Cambridge: Cambridge University Press.

Hubert-Wallander, B., & Boynton, G.M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision* 15(5), 1–12.

Hume, D. (1748). *An Enquiry concerning Human Understanding*. L. A. Selby-Bigge (ed.), 2nd edition (Oxford: Clarendon Press, 1902).

Im, H.Y., & Chong, S.C. (2014). Mean size as a unit of visual working memory. *Perception* 43, 663–676.

Irwin, D.E. & Yeomans, J.M. (1986). Sensory registration and informational persistence. *Journal of Experimental Psychology: Human Perception and Performance* 12(3), 343–360.

Irwin, D.E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 18(2), 307–317.

Irwin, D.E., & Andrews, R.V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J.L. McClelland (eds.), *Attention and Performance XVI: Information Integration in Perception and Communication* (Cambridge, MA: MIT Press), 125–155.

Irwin, D.E., & Gordon, R.D. (1998). Eye movements, attention, and transsaccadic memory. *Visual Cognition* 5, 127–155.

Irwin, D.E., & Thomas, L.E. (2008). Visual sensory memory. In S.J. Luck & A. Hollingworth (eds.), *Visual Memory* (Oxford: OUP), 9–41.

Jackson, F. (1977). *Perception: A Representative Theory*. Cambridge: Cambridge University Press.

Jespersen, B. (2012). Recent work on structured meaning and propositional unity. *Philosophy Compass* 7(9), 620–630.

Johnson-Laird, P. (2006). *How We Reason*. Oxford: OUP.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D.R. Davis (Eds.), *Varieties of Attention* (Orlando: Academic Press), 29–61.

Kahneman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology* 24, 175–219.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* 17(11), 4302–4311.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The Case for Mental Imagery*. Oxford: OUP.

Kosslyn, S.M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S.M. (1994). *Image and Brain*. Cambridge, MA: MIT Press.

Kosslyn, S.M. (1995). Mental imagery. In *An Invitation to Cognitive Science: Visual Cognition, Volume 2*, eds. D. Osherson and S. Kosslyn (Cambridge, MA: MIT Press), 267–296.

Kosslyn, S.M., Ball, T.M., & Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of imagery scanning. *Journal of Experimental Psychology: Human Perception and Performance* 4, 47–60.

Kravitz, D.J., & Behrmann M. (2011). Space-, object-, and feature-based attention interact to organize visual scenes. *Attention, Perception, & Psychophysics* 73(8), 2434–2447.

Kripke, S. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Kulvicki, J. (2004). Isomorphism in information-carrying systems. *Pacific Philosophical Quarterly* 85(4), 380–395.

Kulvicki, J. (2006). *On Images*. Oxford: OUP.

Kulvicki, J. (2015a). Analog representation and the parts principle. *Review of Philosophy and Psychology* 6(1), 165–180.

Kulvicki, J. (2015b). Maps, pictures, and predication. *Ergo* 2(7), doi: 10.3998/ergo.12405314.0002.007.

Laurence, S., & Margolis, E. (2012). The scope of the conceptual. In E. Margolis, R. Samuels, and S. Stich (eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (Oxford: OUP), 291–317.

Levin, D.T., & Banaji, M.R. (2006). Distortions in the perceived lightness of faces: The role of race categories. *Journal of Experimental Psychology: General* 135(4), 501–512.

Lewis, D. (1971). Analog and digital. *Noûs* 5, 321–327.

Lewis, D. (1982). Logic for equivocators. *Noûs* 16, 431–441.

Lewis, D. (1997). Finkish dispositions. *The Philosophical Quarterly* 47(187), 143–158.

Lewis, K.J.S., Borst, G., & Kosslyn, S.M. (2011). Integrating visual mental images and visual percepts: New evidence for depictive representations. *Psychological Research* 75, 259–271.

Lewis, M.B., & Edmonds, A.J. (2003). Face detection: Mapping human performance. *Perception* 32, 903–920.

Lin, C., & Fan, K. (2001). Triangle-based approach to the detection of human face. *Pattern Recognition* 34, 1271–1284.

Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Research* 49(10), 1194–1204.

Little, A.C., Jones, B.C., & DeBruine, L.M. (2011). The many faces of research on face perception. *Philosophical Transactions of the Royal Society B* 366, 1634–1637.

Liverence, B.M., & Scholl, B.J. (2011). Selective attention warps spatial representation: Parallel but opposing effects on attended versus inhibited objects. *Psychological Science* 22(12), 1600–1608.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 309, 279–281.

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology* 6(4), 547–569.

Lupyan, G. (2016). Not even wrong: The "it's just X" fallacy. *Behavioral and Brain Sciences* 39, 40–41.

Lycan, W.G. (1987). *Consciousness*. Cambridge, MA: MIT Press.

MacLennan, B.J. (2012). Analog computation. In R. Meyers (ed.), *Computational Complexity* (New York: Springer), 161–184.

Macpherson, F. (2012). Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research* 84(1), 24–62.

Mandelbaum, E. (2010). *The Architecture of Belief: An Essay on the Unbearable Automaticity of Believing*. PhD dissertation, University of North Carolina, Chapel Hill.

Mandelbaum, E. (2013). Numerical architecture. *Topics in Cognitive Science* 5(1), 367–386.

Mandelbaum, E. (2014). Thinking is believing. *Inquiry* 57(1), 55–96.

Mandelbaum, E. (2016). Attitude, inference, association: On the propositional structure of implicit bias. *Noûs* 50(3), 629–658.

Mandelbaum, E. (forthcoming). Seeing and conceptualizing: Modularity and the shallow contents of perception. *Philosophy and Phenomenological Research*.

Mandelbaum, E., & Quilty-Dunn, J. (2015). Believing without reason, or: Why liberals shouldn't watch Fox News. *Harvard Review of Philosophy* 22, 42–52.

Mandik, P. (2012). Color-consciousness conceptualism. *Consciousness and Cognition* 21, 617–631.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, MA: MIT Press.

Martin, M.G.F. (2006). On being alienated. In *Perceptual Experience*, eds. T. Gendler and J. Hawthorne (Oxford: OUP), 354–410.

Matthen, M. (2005). *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception*. Oxford: OUP.

McDowell, J. (1994). *Mind and World*. Cambridge, MA: Harvard University Press.

McKelvie, S.J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. *Journal of Mental Imagery* 19(3-4), 1–106.

McNamara, T.P. (2005). *Semantic Priming: Perspectives from Memory and Word Recognition*. New York: Psychology Press.

Meehan, D. (2002). Qualitative character and sensory representation. *Consciousness and Cognition* 11(4), 630–641.

Meyer, D.E. & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* 90, 227–234.

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., & Hecht, B. (2016). *Proceedings of the Tenth International AAAI Conference on Web and Social Media*, 259–268.

Mitroff, S.R., & Alvarez, G.A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review* 14(6), 1199–1204.

Mitroff, S.R., Scholl, B.J., & Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition* 96, 67–92.

Mole, C. (2011). *Attention is Cognitive Unison*. Oxford: OUP.

Mole, C. (2015). Attention and cognitive penetration. In J. Zeimbekis & A. Raftopoulos (eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives* (Oxford: OUP), 218–238.

Nanay, B. (2010). Attention and perceptual content. *Analysis* 70, 263–270.

Nanay, B. (2013). *Between Perception and Action*. Oxford: OUP.

Nanay, B. (forthcoming). *Seeing Things You Don't See*. Oxford: OUP.

Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Noe, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

Noe, A. (2015). Concept pluralism, direct perception, and the fragility of presence. *Open MIND*.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition* 107, 179–217.

Palmer, S. (1999). *Vision Science: From Photons to Phenomenology*. Cambridge, MA: MIT Press.

Peacocke, C. (1992). *A Study of Concepts*. Cambridge, MA: MIT Press.

Peacocke, C. (2001). Does perception have a nonconceptual content? *The Journal of Philosophy* 98(5), 239–264.

Peacocke, C. (2004). Interrelations: Concepts, Knowledge, Reference, and Structure. *Mind & Language* 19(1), 85–98.

Pearson, J., & Kosslyn, S.M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *PNAS* 112(33), 10089–10092.

Pearson, J., Naselaris, T., Holmes, E.A., & Kosslyn, S.M. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends in Cognitive Sciences* 19(10), 590–602.

Peirce, C.S. (1894). What is a sign?

Persuh, M., Genzer, B., & Melara, R.D. (2012). Iconic memory requires attention. *Frontiers in Human Neuroscience* 6(126), doi: 10.3389/fnhum.2012.00126.

Pettigrew, R. (2013). Epistemic utility and norms for credences. *Philosophy Compass* 8(10), 897–908.

Phillips, I. (2011). Attention and iconic memory. In C. Mole, D. Smithies, & W. Wu (eds.), *Attention: Philosophical and Psychological Essays* (Oxford: OUP), 202–225.

Pietroski, P. (2008). Think of the children. *Australasian Journal of Philosophy* 86(4), 657–669.

Pinker, S. (1997). *How the Mind Works*. New York: Norton.

Pollatsek, A., Rayner, K., & Collins, W.E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General* 113(3), 426–442.

Potter, M.C., Wyble, B., Hagmann, C.E., & McCourt, E.S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics* 76(2), 270–279.

Prinz, J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.

Prinz, J. (2006). Is the mind really modular? In *Contemporary Debates in Cognitive Science*, ed. R. Stainton (Malden: Blackwell), 22–36.

Prinz, J. (2011). Has mentalese earned its keep? On Jerry Fodor's *LOT 2*. *Mind* 120(478), 485–501.

Prinz, J. (2012). *The Conscious Brain: How Attention Engenders Experience*. Oxford: OUP.

Prinz, J. (unpublished). Faculty psychology without modularity.

Proffitt, D.R. (2009). Affordances matter in geographical slant perception: Reply to Durgin, Baird, Greenburg, Russell, Shaughnessy, and Waymouth. *Psychonomic Bulletin & Review* 16(5), 970–972.

Pylyshyn, Z. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin* 80(1), 1–24.

Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition* 32(1), 65–97.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22(3), 341–365.

Pylyshyn, Z. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences* 25, 157–238.

Pylyshyn, Z. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.

Pylyshyn, Z. (2007). *Things and Places: How the Mind Connects with the World*. Cambridge, MA: MIT Press.

Pylyshyn, Z. (2008). The empirical case for bare demonstratives in vision. In C. Viger and R. J. Stainton (Eds.), *Compositionality, Context, and Semantic Values: Essays in Honour of Ernie Lepore* (Springer), 255–274.

Pylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision* 3(3), 179–197.

Quilty-Dunn, J. (2015). Believing in perceiving: Known illusions and the classical dual-component theory. *Pacific Philosophical Quarterly* 96, 550–575.

Quilty-Dunn, J. (2016). Iconicity and the format of perception. *Journal of Consciousness Studies* 23(3–4), 255–263.

Quilty-Dunn, J., & Mandelbaum, E. (unpublished a). Against dispositionalism: Belief in cognitive science.

Quilty-Dunn, J., & Mandelbaum, E. (unpublished b). Inferential transitions.

Quine, W.V. (1960). *Word and Object*. Cambridge, MA: MIT Press.

Raftopoulos, A. (2014). The cognitive impenetrability of the content of early vision is a necessary and sufficient condition for purely nonconceptual content. *Philosophical Psychology* 27(5), 601–620.

Ramachandran, V.S. (1988). Perceiving shape from shading. *Scientific American* 256(8), 76–83.

Rayner, K., McConkie, G.W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology* 12, 206–226.

Reid, T. (1764). *Inquiry into the Human Mind on the Principles of Common Sense*. D.R. Brookes (ed.), University Park, PA: Pennsylvania State University Press, 1997.

Rescorla, M. (2009a). Cognitive maps and the language of thought. *British Journal for the Philosophy of Science* 60(2), 377–407.

Rescorla, M. (2009b). Predication and cartographic representation. *Synthese* 169, 175–200.

Rescorla, M. (2014). The causal relevance of content to computation. *Philosophy and Phenomenological Research* 88, 173–208.

Reverberi, C., Pischedda, D., Burigo, M., & Cherubini, P. (2012). Deduction without awareness. *Acta Psychologica* 139, 244–253.

Rock, I., ed. (1997). *Indirect Perception*. Cambridge, MA: MIT Press.

Rosenthal, D. (2005). *Consciousness and Mind*. Oxford: OUP.

Rosenthal, D. (2010). How to think about mental qualities. *Philosophical Issues* 20(1): 368–393.

Rumelhart, D.E., & McClelland, J.L. (1986). *Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations*. Cambridge, MA: MIT Press.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Schneider, S. (2009). The nature of symbols in the language of thought. *Mind & Language* 24(5), 523–553.

Scholl, B.J. (1997). Reasoning, rationality, and architectural resolution. *Philosophical Psychology* 10(4), 451–470.

Scholl, B.J., & Leslie, A.M. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In E. Lepore and Z. Pylyshyn (eds.), *What is Cognitive Science?* (Oxford: Blackwell), 26–73.

Scholl, B.J., Pylyshyn, Z.W., & Franconeri, S. (1999). When are spatiotemporal and featural properties encoded as a result of attentional allocation? *Investigative Ophthalmology and Visual Science* 40(4), S797.

Schut, M.J., Fabius, J.H., Van der Stoep, N., & Van der Stigchel, S. (2017). Object files across eye movements: Previous fixations affect the latencies of corrective saccades. *Attention, Perception, & Psychophysics* 79, 138–153.

Schwitzgebel, E. (2013). A dispositional approach to attitudes: Thinking outside of the belief box. In N. Nottelman (ed.), *New Essays on Belief: Constitution, Content, and Structure* (New York: Palgrave Macmillan), 75–99.

Searle, J.R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge, MA: Cambridge University Press.

Searle, J.R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.

Segal, G. (2000). *A Slim Book about Narrow Content*. Cambridge, MA: MIT Press.

Sellars, W. (1956). *Empiricism and the Philosophy of Mind*. 1997 edition. Cambridge, MA: Harvard University Press.

Sergent, J., & Signoret, J.L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex* 2(5), 376–388.

Shea, N. (2011). New concepts can be learned. *Biology and Philosophy* 26, 95–110.

Shea, N. (2014). Distinguishing top-down from bottom-up effects. In D. Stokes, M. Matthen, & D. Biggs (eds.), *Perception and Its Modalities* (Oxford: OUP), 73–91.

Shepard, R.N. (1978). The mental image. *American Psychologist* 33(2), 125–137.

Shepard, R.N., & Cooper, L.A. (1982). *Mental Images and Their Transformations*. Cambridge, MA: MIT Press.

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703.

Siegel, S., & Byrne, A. (forthcoming). Rich or thin? In B. Nanay (ed.), *Current Controversies in Philosophy of Perception* (Routledge).

Sligte, I.G., Scholte, H.S., & Lamme, V.A.F. (2008). Are there multiple visual short-term memory stores? *PloS One* 3(2), e1699.

Smith, A.D. (2002). *The Problem of Perception*. Cambridge, MA: Harvard University Press.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence* 46, 159–216.

Soteriou, M. (2000). The particularity of visual perception. *European Journal of Philosophy* 8(2), 172–189.

Speaks, J. (2005). Is there a problem about nonconceptual content? *Philosophical Review* 114(3), 359–398.

Spelke, E. (1988). Where perceiving ends and thinking begins: The apprehension of objects in infancy. In A. Yonas (ed.), *Perceptual Development in Infancy: Minnesota Symposium on Child Psychology, Vol. 20* (Hillsdale, NJ: Lawrence Erlbaum), 197–233.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied* 74(11), 1–29.

Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.

Stalnaker, R. (1998). What might nonconceptual content be? *Philosophical Issues* 9, 339–352.

Stevenson, R.J., & Boakes, R.A. (2003). A mnemonic theory of odor perception. *Psychological Review* 110, 340–364.

Stich, S. (1985). *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, MA: MIT Press

Stokes, D. (2014). Cognitive penetration and the perception of art. *Dialectica* 68(1), 1–34.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.

Suchow, J.W., Fougnie, D., Brady, T.F., & Alvarez, G.A. (2014). Terms of the debate on the format and structure of visual memory. *Attention, Perception, & Psychophysics* 76(7), 2071–2079.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience* 1(3), 183–184.

Susilo, T., McKone, E., & Edwards, M. (2010). Solving the upside-down puzzle: Why do upright and inverted face aftereffects look alike. *Journal of Vision* 10(13), 1–16.

Tartaglia, E.M., Bamert, L., Mast, F.W., & Herzog, M.H. (2009). Human perceptual learning by mental imagery. *Current Biology* 19, 2081–2085.

Triesman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology* 12, 97–136.

Tye, M. (1991). *The Imagery Debate*. Cambridge, MA: MIT Press.

Tye, M. (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.

Tye, M. (2006). Nonconceptual content, richness, and fineness of grain. In T. Gendler and J. Hawthorne (eds.), *Perceptual Experience* (Oxford: OUP), 504–530.

Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.

Ullman, S. (1996). *High-level Vision*. Cambridge, MA: MIT Press.

Vickery, T.J., & Chun, M.M. (2010). Object-based warping: An illusory distortion of space within objects. *Psychological Science* 21, 1759–1764.

White, T.L. (1998). Olfactory memory: The long and short of it. *Chemical Senses* 23, 433–441.

Wittgenstein, L. (1953). Philosophical investigations. G. E. M. Anscombe (tr.), second edition, 1958. Oxford: Basil Blackwell.

Wu, W. (2014). *Attention*. New York: Routledge.

Zhou, K., Luo, H., Zhou, T., Zhuo, Y., & Chen, L. (2010). Topological change disturbs object continuity in attentive tracking. *PNAS* 107(50), 21920–21924.