ABSTRACT

Title of Dissertation:	SUBJECTIVITY AND COGNITIVE SCIENCE: A DEFENSE OF THE REPRESENTATIONAL THEORY OF CONSCIOUSNESS
	Christopher F. Masciari, Doctor of Philosophy, 2023
Dissertation directed by:	Distinguished University Professor Peter Carruthers, Department of Philosophy

This dissertation provides a defense of reductive representationalism about consciousness. After an introductory chapter, chapter 2 provides a representationalist account of olfaction. In the literature, Burge's (2010) account of representation is widely endorsed. According to his account, perceptual representation represents "objectually", that is, it represents features of the world, as objective. This depends on perceptual constancies. Many authors attempt to defend representationalism about olfaction by showing that there are olfactory constancies. I argue that there are none. Instead, I show that representationalism regarding olfaction is correct by showing that olfaction represents minimally. I then argue that representations in Burge's sense are constructed when minimal olfactory content is embedded in object-files that contain other non-olfactory properties that meet Burge's criteria for representation. In chapter 3, I defend a particular reductive representationalist account of consciousness—the global workspace

theory—against an alternative which suggests that consciousness is richer than the global workspace theory claims. I argue that experience is richer than is standardly suggested by proponents of the global workspace theory, but less rich than the alternative theory suggests. I argue that there are additional resources available to defenders of the global workspace theory in accommodating intuitions of richness that have yet to be fully appreciated by participants in the debate. In chapter 4, I defend reductive representationalism against a new objection presented by Adam Pautz (20172020). He recently suggested that there are several constraints on experience, known as "The Laws of Appearance," that put pressure on the representationalist thesis about conscious experience because they suggest that experience is constrained in ways that representations are not. Since the representationalist claims that experience just is a matter of representing the world to be a certain way, the representationalist owes us an explanation, or else representationalism is false. I argue that the laws are not genuine laws, but that we have the intuition that they are because of the limits of imagination. As a consequence, I show that representationalism is not threatened.

SUBJECTIVITY AND COGNITIVE SCIENCE:
A DEFENSE OF THE REPRESENTATIONAL THEORY OF CONSCIOUSNESS

by

Christopher F. Masciari

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Peter Carruthers, Chair
Professor Georges Rey
Professor Harjit Bhogal
Professor Elizabeth Schechter
Professor Yi Ting Huang

Dedication

For my parents and brother.

Acknowledgements

I am immensely grateful for all the people that have supported me throughout the writing of this dissertation and prior. I would like to begin by acknowledging the integral role my lovely family has played in shaping me as a philosopher. For years they put up with, and often encouraged, my constant need for argument. In particular, I'd like to thank my mother, Andrea Serra Masciari, and my father, Vincent Masciari for their unwavering reassurance that I could do whatever I want in life.

I would also like to thank Peter Carruthers for being the best advisor I could ask for. His endless patience and kindness were crucial to my success during my PhD. His dedication as an advisor extends far beyond the norm and his profound insight helped mold me as a philosopher and as a professional more generally. I am deeply indebted to him.

I owe many thanks to Georges Rey as well. His endless capacity to provides deep and insightful feedback helped shape my work substantially. I am also extremely grateful for our long lunches spent arguing about philosophy. Never has four hours seemed so short.

Many thanks are also owed to my others committee members Elizabeth Schechter, Yi Ting Huang, and especially Harjit Bhogal, who not only gave me insightful feedback on my work but also gave me valuable advice on how to succeed as a graduate student.

At the University of Maryland, College Park, I was fortunate to find an engaging and nurturing intellectual community which helped cultivate many of my

philosophical interests. Special thanks to Sam Kerstein for helping develop my interest in biomedical ethics and for inviting me to join the ethics committee at Capital Region Medical Center. Special thanks as well to Patricia Scripko who was a significant help in developing my values for helping others through biomedical ethics.

Thanks also to Julius Schöenherr, Aida Roige Mas, Andrew Fyfe, Samuel Warren, Shen Pan, and Heather Adair, whom I'm lucky to call friends. Our endless conversations helped shape me into the philosopher and person I am today.

I am also grateful to many professors and colleagues at the University of Maryland, and elsewhere, including, Louise Gilman, Lindley Darden, Rachel Singpurwalla, Dan Moller, Eric Pacuit, Michael McCourt, Aylon Manor, Michael Good, Chiara Palazzolo, and many others.

I would also like to thank my professors and colleagues at San Francisco State University. In particular, I would like to thank Carlos Montemayor, my MA advisor, for all of his countless advice and help towards my current successes. I would also like to thank my friends from SFSU: Tyler Olsson, Francisco Camacho, Patrick Skeels, Matthew Lawrence, and Varia Garon.

Special thanks also to my undergraduate professors Steven Levine and Jennifer Radden. If it weren't for the both of them I would not be the philosopher I am today. Additionally, I would like to thank my friends from UMB: Nick Brotschi, Ted Freedman, and Jake Stanish.

My best friend Joey Jebari also deserves special thanks. We developed as philosophers together, spending hours upon hours over the course of many years

arguing about pretty much every area of philosophy. I am truly fortunate to call him a friend.

I would finally like to thank my partner Lia Curtis Fine for her warmth and encouragement as well as the laughter we shared. All of which got me through the last months of my dissertation.

Table of Contents

Chapter 1: Introduction

> There is a tendency to seek an objective account of everything before admitting its reality. But often what appears to a more subjective point of view cannot be accounted for in this way. So either the objective conception of the world is incomplete, or the subjective involves illusions that should be rejected. -- Thomas Nagel, Mortal Questions, p.196

1. The background

The topic that I aim to address in my dissertation is nicely summarized by the epigraph above. In Thomas Nagel's, "Mortal Questions," he discusses the interplay between the objective realm of third-personal or impersonal facts, and the first-personal, perspective-dependent phenomena that govern many of our intuitions about what the world is like. His discussion is quite broad and is supposed to apply to questions about ethics, metaphysics, broadly construed, and minds in particular.

The so-called "hard problem" of consciousness is perhaps the more perspicuous example of the tension between the first-person perspective and the third-person perspective. The problem is roughly of how to explain the fact that some physical properties, processes, or states are conscious, whereas some are not. According to physicalism, which most take to be a very plausible metaphysical theory, the world is constituted solely by physical properties, processes, and states, yet some of those properties, processes, or states, feel like something have, or to undergo, or to be in. My brain is composed solely of physical stuff, yet I am not an unfeeling object. There are a number of ways to express this concern, however, though all attempt to show that there is an epistemic, and consequently, a

metaphysical, disconnect between the way we understand nature (as being essentially physical), and the way we understand our own experiences.

Perhaps the most notorious way of framing the problem is to point out there is an apparent explanatory gap between the physical idiom on the one hand, and the way we think about phenomenal consciousness on the other (Levine, 1983). Roughly, the idea is that giving an account of the mind and its processes in terms that are consonant with physicalism does not *explain* why there should be any corresponding subjective "feels" (i.e., qualia). This observation has been made forcefully by Jackson (1982) and Chalmers (1996) as well. I will briefly discuss the various ways in which philosophers have defended this observation, but it is important to note that they all have a similar character. Specifically, they all move from epistemic considerations to metaphysical conclusions. In each of the cases to be discussed, failure to a priori identify the physical with the phenomenal, or to *see how* an account of the physical is also an account of the phenomenal, is taken to show that there is a metaphysical distinction between the physical and the phenomenal such that there are separate domains of facts which, if they are related at all, are related merely contingently.

Frank Jackson (1982) has famously suggested that there is an epistemic gap, and thus a metaphysical gap, between physical facts and phenomenal facts. He invites us to imagine a neuroscientist who knows all the facts about the color. She also knows everything there is to know about the human brain and its relation to whatever properties ultimately explain, say, redness. That is, she knows everything there is to know about the perception of redness understood in a third-personal way. Mary is, however, trapped inside a black and white room and so has never seen red herself.

Once she is released from her shackles and sees her first ripe tomato, she is supposed to learn something new. The intuition that we are supposed to have is that there is a fact that Mary does not know before her release, namely, what it is like to experience red. Jackson's contention is that if this is true, then physicalism is false. (Though see Jackson, 2004 for a change of view).

Chalmers (1996) notoriously argued for an explanatory gap by purporting to show the metaphysical possibility of philosophical "zombies". These creatures are functional and physical duplicates of ordinary humans but lack an inner light; they lack phenomenal consciousness. The argument also begins by asserting a conceptual possibility (that zombies are conceivable) and concludes with a metaphysical possibility (that there is world that is a physical duplicate but without consciousness). The idea is that our failure to a priori identify the physical/functional properties uncovered by science with the properties of which we are introspectively aware entails that they enjoy an independent existence.

2. What is consciousness?

The term "conscious experience", or "consciousness" from here on out, is polysemous in the philosophical and psychological literature. (See Block, 1995, Chalmers, 1996, Rosenthal, 2005, Carruthers, 2019, Masciari and Carruthers, 2021). I want to quickly clear the air of any confusions and specify how I am using the term, though there can be some innocent equivocations between the various notions without much loss.

"Consciousness" is sometimes used to pick out a property of an organism as a whole, and other times it is used to pick out a property of a particular mental state of an organism. The former is typically known as "creature consciousness" whereas the latter is known as "state consciousness". Both terms are themselves polysemous. "Creature consciousness" might be understood as specifying the property of being awake, as opposed to being asleep; sometimes also known as 'intransitive creature consciousness' (Kriegel, 2003). Another possible meaning of "conscious experience" is transitive creature consciousness (Rosenthal, 2005). It is often invoked in situations in which some organism is said to be conscious of a property or entity in the world.

In contrast, state consciousness is applied to mental states, not organisms as a whole. State consciousness could also mean a couple of distinct things, however. One might intend to use the term to refer to "access consciousness", to "phenomenal consciousness", or to "higher-order consciousness". Access consciousness is the property of a mental state's being available for reflection, reporting, decision making, or planning. Typically, the contents of access conscious are taken to be coextensive with the properties that are encoded in working memory or short-term memory. Phenomenal consciousness (or a phenomenal property), on the other hand, is the property of there being something it is like to see red, for example. Phenomenal consciousness is a property of a mental state that gives it a first-person accessible feel or "what-it-is-likeness" (for example, the feel of experiencing red) (Chalmers, 1995). In addition, "higher-order consciousness", the property discussed by higher-order theories, is the property of being in a state that one is conscious of having. This

dissertation is specifically about puzzling phenomenally conscious mental states that can be thought to raise a challenge to physicalism.

3. Reducing the phenomenal?

To a first approximation, closing the gap between the physical and the phenomenal would require finding a brain process that is, at a minimum, systematically coincidental with phenomenal consciousness. Often people put this in terms of an attempt to uncover the neural *correlates* of consciousness. Of course, further argument is required to show that one reduces to the other. For instance, one might need to analyze the concept of consciousness in terms that do not make use of phenomenal notions, or at least show that a failure to reduce one concept to the other does not entail that they are *metaphysically* distinct. But establishing that there are correlates of consciousness would be a promising first step towards defending a physicalist account of phenomenal consciousness. At the end of this introduction I provide a standard way of explaining away the intuition that the failure to analyze the concept of consciousness entails that consciousness is not physical.

There are many distinct neurological accounts on consciousness in the literature, but I focus on the Global Workspace Theory in this introduction since it is the operative account throughout much of this dissertation. Originally formulated by Bernard Baars (1988) as a cognitive theory of access consciousness, it has recently been given a neuronal foundation (Dehaene, 2007, 2014). According to Dehaene (2014), "[w]hen we say that we are aware of a certain piece of information, what we mean is this: the information has entered into a specific storage area that makes it

available to the rest of the brain. Among the millions of mental representations that constantly crisscross our brains in an unconscious manner, one is selected because of its relevance to our present goals. Consciousness makes it globally available to all our high-level decision systems. We possess a mental router, an evolved architecture for extracting relevant information and dispatching it." (Dehaene, 2014, 163)

We can think of the workspace as the storage space of working memory. There is some evidence that supposedly shows that this information need not be encoded in working memory, but rather simply *broadcast* to various subsystems, but these debates do not concern me here. (See Carruthers, 2015, 2017 for a discussion). At a minimum the theory is committed to the idea that the information that is broadcast to various subsystems for use in decision making, planning, reporting, reflecting, and so on, that is access conscious. Importantly, the mechanism that is required for informational broadcast is top-down attention. There is an abundance of evidence supporting the idea that attention is necessary for consciousness. For instance, studies on *inattentional blindness* have shown that salient items often go unnoticed when people's attentional resources are exhausted. (Mack and Rock, 1998; Simons and Chabris, 1999). Famously, people failed to notice the presence of a man in a gorilla suit when they were tasked with counting the number of passes of a basketball.

Importantly, however, there is widespread agreement that access conscious states are at a minimum, nomically co-extensive with phenomenally conscious states. Block (1995, 2011, 2014) and Bronfman et al., (2014) are prominent exceptions to

this view but I won't discuss their views here as the third chapter is a direct response to their arguments.

For our current purposes it is important to point out that given the theory's commitment to representationalism, and the evidence that access conscious states are co-extensive with phenomenal consciousness, we are part of the way there to providing a fully physicalist account of phenomenal consciousness.

Now, central to this theory is the concept of representation since it is the information bearing states that play the relevant sort of cognitive role to enable consciousness. It should be noted that the commitment to a representational theory of consciousness predates any efforts to reduce consciousness to states of the brain, especially to its information bearing states. Indeed, part of the reason the representational theory of perception gained in popularity was simply as a response to many of the metaphysical and epistemic problems related to the standard sense data view of the early 20th century. For instance, commitment to a representational theory of mind allowed one to avoid positing obscure metaphysical entities, but it also allowed one to avoid skeptical worries that arise as a consequence of positing intermediaries between minds and the world. Whether or not representationalism does make good on the latter point is orthogonal to the current dissertation, so I ignore it.

An additional motivation for taking representationalism seriously was the idea that perceptual states are intuitively intentional, or about the world. The famous transparency argument makes this intuition clear. The idea is that when one reflects on their experience, it is the world that is presented, not any intrinsic features of the experience itself. So, independent of any commitment to the information bearing

states of the brain, representation was intuitively plausible. This latter point will be central to the discussion below when I discuss objections to representationalism. It should be noted, however, that representationalism *itself* is not committed to the transparency argument, but only externalist versions of representationalism. (see Rey, 1998, and Aydede, 2019, for internalist conceptions of qualitative experience which reject the transparency datum).

But perhaps more importantly, speaking in terms of representation has been standard in cognitive science since the early 1960's, and since then people have been committed to realism about representation in virtue of the widespread success of cognitive science explanations. There are some defectors, as I mention below, but they are decidedly outside of the mainstream of cognitive science and do not have obvious explanatory merits.

4. What is representationalism?

Like the concept of consciousness, "representationalism" is likewise a multifarious notion. Minimally, any theory that is representationalist claims that experience necessarily has a representational component. This understanding of an experience being representational just amounts to the claim that representation is a necessary condition of having an experience. For an experience to have a representational component is just for it to partly be about some state of affairs, real or unreal. So, my experience of a cat being on my lap represents that "there is a cat on my lap". Crucially, an experience of a cat being on my lap is veridical if and only if there is a cat on my lap, and hallucinatory or illusory if there is not a cat on my lap.

This is known as *weak representationalism*, and it is the least radical view one could hold regarding the necessity of mental representation for having an experience. Weak representationalists simply claim that all experiences are *partly* constituted by representational states. According to the weak representationalist, every experience represents the world being a certain way (Chalmers, 2004; Crane, 2003; Levine, 2003; Lycan, 2001) It is a decidedly weak thesis, and it is compatible with experience only being *fully* explained by additional metaphysical properties or entities such as qualia or sense-data. Moreover it is largely undisputed (though see Gibson, 1979; Chemero, 2001; Noe, 2004; Shapiro, 2007 for arguments against this idea).

Pure representationalism is the most radical view one could hold regarding the sufficiency of mental representation for having an experience (see Chalmers, 2004). It is the claim that if a mental state is representational, then it is necessarily phenomenally conscious. I won't spend any time going into any of the views defending this idea since it is a nonstarter; most everyone agrees that there are unconscious representational states; in particular it flies in the face of the mainstream of cognitive science. I only mention this option since it is logically possible.

Lastly, strong representationalism, which will be the focus of this dissertation, is the claim that phenomenal consciousness is fully explained by a certain kind mental representation with a certain *functional role*. There are various debates about exactly what the functional role must be, however. For instance, many claim that mental representations need to be accessible for use in reasoning or planning (Dehaene, 2014, Carruthers, 2019, Masciari, 2021). That is, in order for a representation to count as phenomenally conscious, at a minimum, it needs to be

access conscious, though some argue that this isn't sufficient since they also claim that the contents need to also be *nonconceptual*. Alternatively, higher-order theories claim that representations are conscious if (1) one has immediate higher-order thoughts about them (e.g., I'm currently perceiving a red tomato), (2) one is disposed to have higher-order thoughts about them, or (3) one "perceives" (with an *inner sense*) that one is currently in a mental state that represents some property or whatever. (See Gennaro, 2018, for an overview of the various options). What is common among all these distinct views is that in order for a mental representation to count as phenomenally conscious it needs to play the right sort of functional role. This dissertation takes strong representationalism and the Global Workspace Theory for granted as a starting place, but I defend them against objections in the following chapters. In virtue of the fact that my main focus is on reconciling our first-personal access to our minds with the best cognitive science explanations on offer, I largely ignore both pure and weak representationalism. All mentions of the term "representationalism," unless otherwise indicated, are meant to refer to strong representationalism.

5. Some objections to representationalism

Before I discuss my plan in this dissertation, I will briefly survey objections to the representationalist account of consciousness. The objections apply to strong representationalism independently of whether or not representationalism is defended as a physicalist thesis (Chalmers, 2004). So if one thought, for instance, that the Global Workspace theory is the correlate of consciousness, and one defended strong

representationalism against objections, it would still be a logical possibility to deny physicalism. Indeed, this is exactly the argumentative strategy implicit in the zombie thought experiment.

In any case, the objections I present below are all supposed to show that there could be phenomenal differences without any representational differences. If any of the cases are genuine, that would count against representationalism (of the strong variety) since they would contradict the claim that representation (of a certain kind and which plays a certain role) is sufficient for phenomenal consciousness. Call this set of objections phenomenal-difference-representational-sameness objections.

Peacocke (1983) suggests that there is a phenomenal difference between looking at the layout of one's room with both eyes open and with one eye closed, yet the very same layout is represented. Boghossian and Velleman (1989) argue that there is a phenomenal difference between looking at an image with one's glasses on and with one's glasses off. Yet, there is not a representational difference; in both cases I see a cat, but in one case I see it clearly, and in the other I see it blurrily.[1] MacPherson (2006) suggests that there is a phenomenal difference between seeing a diamond

[1] Here's what they say specifically:
"There are other, more familiar cases that refute intentionalism in a similar way. These, too, are cases in which something is seen without being represented in the content of experience as intentionalism would require. If you press the side of one eyeball, you can see this line of type twice without seeing the page as bearing two identical lines of type. Indeed, you cannot even force the resulting experience into representing the existence of two lines, even if you try. Similarly, you can see nearby objects double by focusing on distant objects behind them, and yet you cannot get yourself to see the number of nearby objects as doubling. And by unfocusing your eyes, you can see objects blurrily without being able to see them as being blurry. None of these experiences can be adequately described solely in terms of their intentional content. Their description requires reference to areas of colour in a visual field, areas that split in two or become blurry without anything's being represented to you as doing so." (p.94) And in the next paragraph, "We therefore endorse Peacocke's decision to posit a visual field with intrinsic sensational qualities." (p.94)
By "intrinsic sensational qualities", I take them to mean "qualia" in a very narrow sense of the term.

figure as an upright diamond or as a rotated square, yet there is not a representational difference. Block (2010) suggests that if we are presented with two red lights presented on a black background and we shift our attention from one to the other there will be a phenomenal difference (the vividness of the focally attended point of light is greater than the peripherally attended one) without there being a representational difference. All cases are supposed to undermine strong representationalism. In each case the representationalist responds by arguing that even though the representational content *seems* to remain constant, there are subtle representational or informational differences in each case that can be accounted for empirically.

I will briefly mention some plausible (albeit potentially defeasible) responses put forward by representationalists. Regarding the claim that there is a phenomenal difference between seeing binocularly and monocularly, Tye (1992) claims that when one sees binocularly one sees more of the objects and sees them with more determinacy than when one sees monocularly (1992, p.174). Although put slightly cryptically, the claim is that one sees more depth in the former case and less in the latter. This is surely a representational difference since depth seems to be a property of the world, though perhaps largely entangled with one's perspective. Regarding the case of blurry vision, the same sort of response seems to apply. Although one might still represent the abstract content "that there is a cat in front of me" in both cases, when one sees clearly, one is representing the content in a more determinate manner. Perhaps one sees the cat's fur with higher resolution.

These observations are reasons for thinking that phenomenally conscious states are individuated in virtue of their nonconceptual (as opposed to conceptual) content. There has been a huge literature about exactly how to think about the contrast between conceptual and nonconceptual content, however. For example, Dretske (1981) argues that digital (as opposed to analog or nonconceptual) representations do not carry any other information other than what is expressed by the proposition which constitutes the representation. However, when once perceptually represents, say, the presence of a cat, one simultaneously represents its color and size, etc. Evans (1982) argues that we can discriminate more colors than we can name. Peacocke (1986) argues that many perceptions are represented in a "unit-free" way (for instance, the tree is *far from me*, but without any commitment to the appropriate unit. More recently, Beck (2012) argues that analogue magnitude representations (representations that are structurally isomorphic to what they represent) do not satisfy the generality constraint; a constraint that is taken to be central to propositions. (See Bayne and Montague (eds.) 2011 for an extensive overview of the various positions). In much of this dissertation I presuppose the view that phenomenal conscious mental states are those that are nonconceptual (see Carruthers, 2015 for a defense of this position).

The objections put forward by MacPherson and Block can be explained similarly, by appealing to nonconceptual content. Consider the diamond-square shape experience (of a kind with typical Gestalt shift cases such as the notorious duck-rabbit). Although one initial reply put forward by Tye (2000) was that when one sees the shape as an upright diamond, as opposed to a tilted square, the subject employs

distinct concepts, Orlandi (2011) rightfully argues that this line of thought is in tension with the standard commitment to sensory qualities being fully nonconceptual. In virtue of this problem, one might instead represent different possibilities of motion in a fully nonconceptual way; the diamond is easily tippable, whereas the square sits firmly. However, whether or not motor representations are constitutive of *visual* experiences *as such*, or simply constitutive of experience as a whole is beyond the scope of this introduction, so I put this to one side. Lastly, consider Block's case of phenomenal contrast. The reason that the light that one is focally attending to seems brighter and, in some sense, larger than the peripherally attended light, is explained by attention. Block takes this to show that there are phenomenal differences that are explained by attention and not by an experience's representational content. However, one line of reply is to suggest that when one focally attends to one of the dots, they thereby represent it as larger and brighter, albeit nonconceptually. Of course this would count as a misrepresentation, but this is perfectly in line with a representational account of experience. Counterexamples to representationalism that issue from introspection abound, but this brief survey suffices to provide the general character of many of the apparent counterexamples.

6. Physicalist representationalism?

Now, it should be noted that even if one thinks that strong representationalism is true, and one also thinks that the global workspace theory of consciousness (or any other neural correlate account) is true, or at least approximately true, there is still the further question about reduction. As Chalmers points out, it is conceivable that there

could be physical and functional (and also representational duplicates) to us, but that lack phenomenally conscious experiences; there is an explanatory gap. So, even if phenomenal consciousness is correlated with global broadcasting, and strong representationalism is true, the relation between globally broadcast states (understood with respect to representation) and phenomenal consciousness might not be necessary.

A traditional reply to this line of thought is to claim that the explanatory gap (or the conceivability of zombies) is a special case of referential opacity, and thus there is no reason to think that these are genuine possibilities. People point out that are many cases where we do not have a priori access to metaphysically necessary identities, but wherein physicalism is not threatened. For example, I might know that I enjoy drinking water, but I might fail to know that I enjoy drinking H2O. Now, in ordinary cases of *a posteriori* identities (water *is* H20, or heat *is* mean kinetic energy), there is no insoluble *epistemic gap* since we are often able to grasp *how* heat could simply be the movement of molecules or how water is identical to H20. But the question is why, regarding phenomenal consciousness, the epistemic gap remains. That is, why do we find zombies conceivable? What is ordinarily claimed is that there is some sort of conceptual disconnectedness that *explains* why we resist the identification of the physical with the phenomenal even though they are as a matter of metaphysical fact, identical. There has been an abundance of discussion about how to make sense of this conceptual disconnectedness, but for our purposes most phenomenal concepts are typically understood to be inferentially *isolated* from

physical concepts.[2] This conceptual isolation is supposed to explain why we *think* that our phenomenal states are merely contingently related to our brain even though phenomenal consciousness reduces to facts about the brain. This strategy is known as the Phenomenal Concept Strategy. (Though, see Block and Stalnaker, 1999, for a related view that argues that the failure to analyze the concept of consciousness in physical terms does not entail metaphysical dualism). The phenomenal concept strategy is not only utilized as a response to Jackson (1982) but also to more recent examples. It is useful since it allows us to fully characterize phenomenal consciousness without reference to phenomenal properties.

The upshot of the previous discussion is that if one successfully provides the physical correlates of consciousness, such as the Global Workspace Theory, defends strong representationalism against objections, and provides a good reason for why we are compelled by the possibility of zombies (or the explanatory gap in general), one would have a quite good reason to think they've provided a fully physically reductive, representational theory of consciousness. This dissertation attempts to support the first two goals. Chapters 1 and 3 defend strong representationalism against two objections and chapter 2 defends the Global Workspace Theory against an alternative theory endorsed by Block.

[2] Now, although phenomenal concepts are largely isolated from physical concepts, they might not be completely if they serve as evidence in general reasoning. So, if one can invoke a phenomenal concept to explain an action, for example, one might think it is at least partly promiscuous. I leave this possibility to the side in what follows.

7. Map of dissertation

The following is the plan of my dissertation:

In chapter 2, "Stop and Smell the What? Two Kinds of Olfactory Representation", (Masciari, 2022; published in *Synthese*), I attempt to provide an account of olfactory experience that takes our best cognitive science explanations seriously, namely, the computational-representational theory of mind, while also attempting to do justice to the object-oriented phenomenology of olfaction. There are many accounts of representation in the philosophical literature. However, regarding olfaction, Burge's (2010) account is widely endorsed. According to his account, perceptual representation is always as of an objective reality. That is, perception represents objects as objects, where "object" is construed quite broadly to include properties and events. Many authors presuppose this account of representation and attempt to show that the olfactory system produces representations of this sort. I argue that given Burge's account of representation, smell represents objectually, or it doesn't. Some people argue that it does, but I show that their arguments fail. This poses a problem for representationalism since it is clear that we have olfactory phenomenology. Instead I show that we can maintain representationalism by taking seriously a minimal notion of representation, adopted from Nick Shea. I argue that the olfactory system is representational after all even if it doesn't represent objects as objects. Rather, olfaction results in "minimal representations." I nevertheless argue that objectual representations of olfactory objects are constructed when minimal olfactory content is embedded in object-files that contain other nonolfactory properties that meet Burge's criteria for representation. Some extant work on feature

binding, attention, and object-files is introduced to support the suggestion. The final portion of the chapter brings together the technical elements introduced in the earlier parts to suggest a way in which we might better understand olfactory phenomenology.

In chapter 3, "Additional Resources for Sparse Theories of Phenomenal Consciousness" (Masciari, 2021; published in *Journal of Consciousness Studies*), I defend the global-workspace theory of consciousness against an objection. To do so, I engage in the so-called phenomenal overflow debate. This is a debate about the richness of phenomenal consciousness. There are two candidate views: rich and sparse. The rich view says phenomenal consciousness outstrips access consciousness and the contents of working memory. The sparse view, typically defended by supporters of the global-workspace theory of consciousness, denies this. Importantly, according to some conceptions of the sparse view, the intuition we all have, that upon opening our eyes we are confronted with a plethora of sensible qualities, depends partly on perceiving the gist of a scene and partly on underestimating the rate at which our attention shifts; the latter is a phenomenon known as the refrigerator-light illusion. I argue that experience is richer than is standardly suggested by extant cognitive science accounts, but less rich than our phenomenology suggests. In this chapter I show that there are additional resources available to defenders of the global workspace theory in accommodating intuitions of richness that have yet to be fully appreciated by participants in the debate. To this end, research pertaining to feature binding and activity-silent working memory is discussed. The upshot is that many of

our putatively rich experiences can be explained by cognitive mechanisms that are generally accepted.

In the last chapter, “The Bounds of Experience: Representationalism and the Laws of Appearance”, (under review) I defend the representational theory of consciousness against a new and important objection. Adam Pautz (2017/2020) has recently suggested that there are several constraints on experience, known as “The Laws of Appearance”; for instance, that it is impossible to see pure-blue and pure-green at the same location and at the same time. The laws put pressure on the representationalist thesis about conscious experience because they suggest that experience is constrained in a way that Pautz claims representations are not. But since the representationalist claims that experience just is a matter of representing the world to be a certain way, the representationalist owes us an explanation, or else representationalism is false. Importantly, The Laws are derived not from scientific facts about human or animal cognition, but from introspecting on their intuitive plausibility. This sort of dispute confronts our commitments to both phenomenological and cognitive scientific explanations head on. In this chapter, I begin by arguing that the Laws of Appearance only seem necessary. I then consider a view that suggests that if we restrict the scope of experience to normal human perception, we might indeed be justified in our intuitions about the modal strength of the constraints on experience. Ultimately, I argue that either way the Laws of Appearance do not threaten the truth of representationalism understood as a physicalist thesis about experience. This suggests that physicalist varieties of representationalism can account for experience from within a purely third-personal

framework, while also allowing for the truth of our intuitions about our own phenomenology.

8. Why it matters

Although one might have interests regarding each of the chapters of this dissertation, I think that each is actually important for the progress of philosophy and cognitive science generally. The computational-representational theory of mind has seen an explosion of success since its advent in the early 1960's. This is due to a number of reasons. First, it has it provided philosophers with some much-needed relief from various traditional solutions in the philosophy of perception (for example, it did away with the need to posit metaphysically obscure entities such as sense-data). Second, it is also the operative explanatory framework in the mainstream of cognitive science. This latter point is worth emphasizing since, prior to Chomsky's 1959 piece against Skinner, the standard view in psychology was behaviorism (the theory that suggests that all behavior can be explained by conditioning processes without the need to posit any *invisible* mental entities such as beliefs or desires). Although there are some alternative contemporary explanatory streams in cognitive science, such as the embodied cognition paradigm proposed in some versions of ecological psychology, none has seen as much explanatory success as representationalism. I think it is therefore important to ensure that representationalism is on a good footing in all cases if we are to maintain explanatory unity. Of course, one might deny there that there will be a unified theory of mind (and also of phenomenal consciousness) or that we should want one, but I think this should not be accepted unless it is shown that representationalism cannot accommodate all of the

data points. So, although I have not defended representationalism against all plausible objections, I have gone some way to defending it against certain salient and intuitive puzzles. Even if not revolutionary, it is progress, nonetheless.

Chapter 2:
Stop and Smell the What? Two Kinds of Olfactory Representation

1. Introduction

Recently there has been much debate about the nature of olfaction. (Barwich, 2019; Batty, 2010, 2014; Burge, 2009, 2010, 2014; Carvalho, 2014; Millar, 2017; Skrzpulec, 2019.) Some argue that olfaction is not representational[3], others that it is. Interestingly, both those who reject *and* those who accept the representational characterization of olfaction tend to presuppose that sensory representation (as opposed to sensory *information registration*) requires perceptual constancies.[4] That is, capacities to systematically represent stable features of one's environment given variability in proximal input. Importantly, Burge and others think that the content of these representations are "objects", but in a proprietary sense of the term (e.g., color and luminance are both *objects* —from now on we will refer to these as objectual properties to avoid confusion with the commonsense usage. The term "object" will be restricted to ordinary objects such as roses, garlic, coffee, etc.). The appeal to perceptual constancy is supposed to explain the fact that perceptual representation is always of an *object*ive reality, that it takes objectual properties as its target. We will

[3] In what follows, all uses of the term "representation" denote representation of the mental variety; the kind that features in mentalistic explanations. Though, all mental representations are likely ultimately identical to or grounded in mentalistic explanations, this paper makes no commitment to the metaphysics of representation.

[4] Carvalho (2014) is an exception to this trend. He argues that the more minimal notion of figure-ground segregation is sufficient for representation. I will discuss his view in what follows and ultimately claim that it doesn't work.

refer to a representation of this sort as an "objectual representation". As a result, one side argues that there are olfactory objectual representations, and the other argues that there are not.

This article is an attempt at reconciliation. It will be argued that a condition that is both weaker than perceptual constancy, but stronger than mere information registration, is sufficient for representation, or what we will call "minimal representation". Moreover, it will be argued that olfaction satisfies this middle ground position and so issues in minimal representations. However, an argument for how representations of objectual olfactory properties are constructed will also be put forward. The claim to be defended is that our minds construct object-files that contain olfactory content among their total content. This paper will proceed as follows. Section 2 will first survey and evaluate the various arguments in defense of olfactory objectual representation, ultimately showing that they fail. Section 3 sets out the middle ground conception of representation, derived from Shea (2018). Here, it will be argued that olfaction issues in minimal representations. Section 4, however, will apply an account of the constructive nature of perception, derived from standard cases in the visual and auditory domain, to the olfactory domain, thus showing how objectual representations of olfactory objects are constructed. Section 5 will conclude.

2. Olfactory representation

In this section, some of the extant arguments in defense of olfactory objectual representations will be reviewed. There are many views on offer, so only the leading contenders will be considered. As noted above, Burge (2009, 2010, 2014), as well as

proponents of olfactory objectual representations, presuppose that representation of that sort depends on *perceptual constancies.*[5] Perceptual constancies are capacities to represent the very same content (property, feature, relation, etc.) given enormous variations in proximal input. Viewing a circular table from above issues in quite different patterns of light on the retina than viewing it from the side, yet we still represent the very same object, namely, the circular table. Similarly, when viewing the color of a wall under differing lighting conditions, there are changes regarding the frequencies of light impinging on the retina, yet we see that the color has not changed. This process is observed in both vision and audition and is not limited to the human case. For example, there is evidence that chickens also represent perceptual constancies. (Wood, 2013; Wood and Wood, 2018.) Crucially, and this will be important in what is to come, constancy mechanisms are posited to partly explain how creatures solve the so-called "inverse problem". The inverse problem arises from the fact that proximal stimuli underdetermine their causes. That is, proximal inputs to the retina are consistent with a number of interpretations of what the world is like.

Regarding olfaction, some, including Burge, argue that there are no such constancies. They claim that humans, and most non-human animals, are indeed able to *detect* odors, but this capacity is not yet sophisticated enough to properly count as representational.[6] Others argue that there are, in fact, olfactory constancies, and so there are indeed olfactory objectual representations.[7] This article agrees that there are

[5] It should be noted that Burge thinks objectual representations are the only sort of perceptual representation that exist.
[6] By "detect", Burge and others mean mere information registration. Importantly, this notion is non-mentalistic. For example, thermometers can detect temperature, but thermometers do not represent temperature in any sense that matters for psychology.
[7] I provide details of their accounts in what follows.

no olfactory constancies, and thus, agrees that olfaction does not issue in objectual representations. Nevertheless, this article disagrees with the claim that organisms *merely* detect odors. The upshot is that human and non-human animals alike *minimally represent* odors. This claim will be defended in Sect. 3.

Arguments in support of the representational status of olfaction are set out in one of two ways. The arguments either (1) appeal to perceptual constancies (Batty, 2014; Millar, 2017) or (2) they appeal to figure-ground segregation (Carvalho, 2014). It will be argued that only the first way directly confronts Burge's criterion, but that all instances fail, and (2) more or less changes the subject. Both options will be dealt with in turn. Ultimately, they each say something correct about olfaction, but neither meets the conditions of objectual representation. Importantly, they *do* meet the conditions for minimal representation. This will be spelled out in Sect. 3, however.

2.1 Filling-In

Millar (2017) argues that there are two ways in which olfactory constancies can be established, and thus, there are two ways to establish olfactory objectual representations. The first pertains to the "filling-in" of odor compounds, and the second has to do with the perception of intensities. Regarding "filling-in", Millar claims that, "[w]e are able to perceive odours *as the odour-type that they are*, even if they are corrupted or missing elements due to poor perceptual conditions." (p.4300.) The claim is that we can represent the smell of coffee even in the absence of particular odorants (sets of odor molecules).[8] Furthermore, the representation is of the

[8] Millar introduces coffee as a paradigm example of filling-in, so for consistency that will be the example used throughout.

odor *type.* More will be said about her filling-in proposal in what follows. Two objections to this account will be put forward.[9] After that, we will evaluate the argument in favor of alleged olfactory intensity constancy.

Before going over her view, it is worth noting the similarities of her account with the one on offer in this article. According to Millar, figure-ground segregation and intensity constancies can be used to argue that olfaction represents odors as particular, discrete objects (2019, p.4301). The conclusion she reaches shares similarities with the conclusion on offer in this paper. However, it is unclear that the processes she invokes really do apply in the case of olfaction and thus it is unclear that olfaction represents objects on its own. Rather, it seems more plausible that to defend her conclusion, we need to see how olfaction interacts with cognition and other perceptual modalities. This will be taken up in the latter portion of this article.

What is Filling-In?

It is not entirely clear what filling-in in olfaction amounts to. However, Millar claims that we can analogize the case to seeing an object as complete even though it is occluded. Speaking slightly more specifically, Millar claims that,

> Perhaps in the case of olfaction, one could recognise that not all of the properties typical of the coffee odour are apparent at once (picking out, say, the fruitiness and yet needing to move closer to coffee cup to pick out the chocolate notes), and yet still also recognise the coffee odour as complete and unified in another sense. (p.4301)

It seems that the idea of filling-in is that there is a presumed set of features that together, explain the unity of the thing in question, e.g., the cup of coffee. Perhaps the idea is that the coffee is a thing that has distinct properties (distinct

[9] See Barwich (2019; p.6-7) for some interesting empirical reasons to reject Millar's account also.

notes), such that one can identify it without identifying all the constituent parts? On the surface, the visual case and the olfactory case are quite different. There is a sense in which the properties (notes) that any coffee has are arbitrary; there are many varieties of coffee. Shapes, in contrast, are not simply a bundle of arbitrary properties with a presumed unity. However, all shapes have complete sides. So, it may be an assumption built into the visual system that shapes with incomplete appearances really are complete. It is unclear why this should hold for olfaction, however. In any case, supposing the analogy holds, several objections will now be put forward.

Type/Token Objection:

One obvious worry concerning Millar's suggestion is that in vision, perceptual constancies do not lead to the representation of a type, but of a token. When one represents a tree's size given variations in the amount of one's visual field it occupies, one does not, therefore, represent the *type* TREESIZE*,* but rather a particular token of the property, namely its size. When one says, "*that* tree isn't shrinking", one is not thereby committed to the claim that the type is invariant, just that the particular tree is. This is also true with regards to color perception. It is *that* shade of red that is constant, not the type, whatever that could mean. Representing the same type across variations of input, thus seems like a feat of cognition, not perception proper. This point will be elaborated in Sect. 4.

Concept Application or Pattern Matching:

An additional reason to think that olfactory filling-in is not a genuine perceptual constancy is that it depends on concepts. So, to "fill-in", in the way described above, one needs to employ a concept. Although there are plausible reasons

to reject the idea that olfactory filling-in depends on concepts essentially, there is some reason to think it does, at least in some cases. For instance, Herz and Von Clef (2001) presented subjects with various mixtures of odorants, one of which contained the chemical butyric acid. They found that if the butyric acid was labelled with "Parmesan cheese" the subjects rated the smell high on a hedonic scale, whereas if it was presented with the label "Vomit" it was rated low. Presumably this effect depends on the subjects smelling the butyric acid *as cheese* or *as vomit*, thus affecting its quality. Both parmesan cheese and vomit contain many more odorants than butyric acid, but the subjects were seemingly able to "fill-in" the missing information such that they represented the mixture as cheese or as vomit, depending on the concept that was employed. Since the difference in hedonic rating and perceived quality depends on the label, it is plausible that filling-in depends on concepts.

The problem is that if concepts are an important part of the explanation of filling-in, then it is unclear that it is a genuine constancy mechanism since concepts are not part of the explanation of perceptual constancies. It is true that under many, entirely varying, conditions we may utilize the same concept, say, COFFEE, but that is not the notion of a constancy used in psychology, in particular in vision science. Representing that some visual object X is the same X given variations in proximal input does not depend on utilizing a concept. Seeing a Kanizsa triangle as complete does not require the concept TRIANGLE. One need not know the category under which a particular thing falls to see that it is the same thing across variations in input. Indeed, there is a lot of evidence that many non-human animals and human infants

have color constancies without the relevant concepts. (See Cohen, 2015 for a general overview of the psychology of perceptual constancies.)

On the other hand, perhaps a defender of a constancy-based interpretation of filling-in might suggest that instead of employing concepts such as COFFEE, creatures may utilize a sort of non-conceptual pattern matching to determine the odor (perhaps by using something like a prototype). Perhaps the pattern matching occurs within olfaction—there is reason to think this is the case regarding visual filling-in. This would bypass the concept objection. In fact, if this is roughly the correct way to think about olfactory "filling-in", then it does appear to be representational according to Burge's rather strict definition. However, there is some reason to doubt that this is what occurs in olfaction absent concepts or categorization (for example, see Barwich 2019 for detailed empirical arguments against "filling-in" that operates independently of concepts or categorization)[10]. But moreover, even if olfaction utilized some sort of statis- tical pattern matching, it is entirely unclear why, given missing odorants, olfaction would "fill-in" the missing information. The olfactory system could just represent the corrupted compound as what it is: a specific mixture of odorants. Since it is unclear if there are any empirical results that contradict this, it seems more parsimonious to go with the simpler option.

So, it seems there are two options. Since the second option appears to either be empirically inadequate or at least not confirmed, we ought to go with the first option. The claim to be defended is that employing concepts is in fact how we should

[10] Barwich's (2019) primary criticism of filling-in is that odors are not discrete and stable entities. I need not belabor this point here.

under- stand the representation of a stable odor. Crucially, however, this occurs outside of olfaction itself. This point will be elaborated in Sect. 4.

2.2. Intensity

The second type of constancy that Millar (2017) puts forward pertains to an odor's intensity. She claims that "[t]he size of sniff that one takes has a significant effect on the neural response of the olfactory nerve, but does not result in significant changes to the apparent intensity of an odour...We seem able to perceive an invariant feature of an odour—its intensity—even though our perspectives on the world change." (p.4300) Millar likens this to the case where a tree in the distance seems smaller than if you were near to it, yet you still represent the tree as the same size.

Although this is a compelling account of perceptual constancy, it is not without difficulties. First, Barwich (2019) argues that one cannot appeal to the perception of odor intensity to argue for olfactory object representations because intensity is subject to perceptual adaptation. (p. 6–7) That is, although intensity might remain stable for several moments, the longer one is exposed to a smell, the less salient and detectable it will be. So, this criticism essentially amounts to a denial of intensity constancy effects altogether[11].

Second, and perhaps more importantly, Błażej Skrzypulec (2019) argues that Millar's account merely explains the constancy of certain *features*, but not the

[11] Although there is evidence in favor of this hypothesis, it faces problems of its own. It is plausible that there is a brief temporal window in which odor intensity is indeed stable. Even if adaptation occurs, it is arguable that that is insufficient to show that there are no stable representations of odor intensity. The only conclusion that adaptation forces us to reach is that if there are intensity constancies, they are short lived. The following takes no stand on this issue, however. I would like to thank an anonymous reviewer for emphasizing this point.

constancy of the object itself (whatever that may be). His point is that the intensity is merely a *property* of the smell. So, arguing that the feature is constant, does not entail the smell is constant. However, this argument fails for an obvious reason. As noted above, Burge and others mean something quite idiosyncratic by "object". Colors are objects. Shapes are objects. The luminance of a surface is an object. So, arguing that an odor's intensity is a property does not show that it is not an object in the relevant sense. But there is, however, a related objection, and perhaps this is what Skrzypulec is really aiming at. If one is attempting to explain color constancy, the invariance of a color's luminance does not cause one to perceive that the color itself is constant. After all, luminance is just one dimension of color (there is also hue and saturation). Similarly, it is not enough to say that the intensity is the same across qualitatively distinct experiences. So, perhaps the intensity of a smell can indeed be invariant—presumably for a very short time—but the issue is whether the odor itself is. That is, the question is whether the odor quality (rosy, garlicky, etc.) of the smell is invariant.

The arguments cited above are not the only ones that defenders of olfactory objectual properties make. Some, although not all, argue that figure-ground segregation also provides reason to think that olfaction issues in objectual representations. We shall now consider an argument of this sort.

2.3. Figure-ground segregation

Figure-ground segregation is a concept derived from Gestalt psychology. (See Wertheimer, 1923). It refers to a perceptual systems' ability to extract signal from noise (to perceive some stimulus relative to a background). Moreover, figure-ground segregation is integral to visual perception in that it is a *necessary* condition for the

ability to perceive objects in the ordinary sense or objectual properties, in Burge's idiosyncratic sense. For instance, one sees that text is superimposed on a white background, that a person is standing in front of their house, that a red patch is on top of a larger blue patch, and so on. Figure-ground segregation is not limited to the visual case. Indeed, it seems like an ability of any perceptual system. A vocal melody can be heard as distinct the rest of the band, a small bump can be felt as distinct from the page on which it is printed, and a particular smell can be distinguished from a more complex smell in which it is embedded. The question, then, is if figure-ground segregation is *sufficient* for objectual representation. Carvalho (2014) and Wilson & Stevenson (2006) seem to think it is. For instance, Carvalho claims that,

> [o]dors, in this picture, are type-individuated by their chemical structure, and it is the function of the olfactory system to recover these structures out of all the chemical compounds that arrive simultaneously at the nose. The output of this process is a (token) representation of an odor — an olfactory object — that characterizes the content of your olfactory experience. (p.58)

Moreover, Wilson & Stevenson (2006) remark that,

> Odors most likely are never sensed in the absence of background odors, except perhaps within a well-controlled laboratory (even then, it is probably rare). Any given inhalation, therefore, will include odorant features that belong to a target odor and odorant features that belong to the background. The olfactory system seems very capable of filtering out background odors while leaving responsive- ness to novel odors intact. (p. 8)

Although it is central to vision that we can distinguish figures from a background, it is by no means obvious that this is sufficient for objectual representation. Flashes of light, for instance, are not objectual in the relevant sense despite being distinguishable from their background. The reason is that perceiving flashes of light does not involve a perceptual constancy. As Millar (2017) points out, "...flashes aren't perceived as invariant across changes in perspective, but are merely

transient figures that capture our attention." (2017, p.4287.) Furthermore, she claims that after-images are, likewise, distinguished from a background but are also not objectual in the relevant sense. (p.4287–4288.) This all seems correct. Perceiving a color as invariant given changes in proximal input no doubt requires one to see the color as distinct from its background, but that is not enough.

This article is in full agreement that figure-ground segregation is not sufficient for objectual representation. However, as will be argued in Sect. 3, there is an important sense in which figure-ground segregation is implicated in minimal representation. Shea's (2018) account of representation seems equipped to explain this general and striking capacity of perception.

Thus far, it seems that many of the extant defenses of olfactory objectual representations fail. The reason is either because there are, in fact, no olfactory constancies or because figure-ground segregation is simply insufficient to explain objectual representation.[12] This should not lead one to be skeptical of olfactory minimal representations, but rather, just of olfactory objectual representations. As will be spelled out in the next section, an account of minimal representation will accommodate much of what has been said about olfaction without the need to appeal to constancies. As a disclaimer, the account of representation on offer in the next section changes the subject away from Burge's concerns in much the same way that Carvalho's appeal to figure-ground segregation does. However, as will be argued in Sect. 3.2, this changing of concerns is theoretically and practically well-motivated.

[12] See Keller (2016, pp.72–75) for additional reasons to reject the idea that figure-ground segregation is sufficient for the perception of objects.

3. Minimal representation

A notion of minimal representation has been championed by Nicholas Shea in his recent 2018 book, "Representation in Cognitive Science." It will be argued that much of what has been said in support of (alleged) olfactory objectual representations is explicable in those terms, and therefore, grounds olfactory minimal representation.

Shea's (2018) theory of representation is in the lineage of both information-theoretic views of representation (Dretske, 1988) as well as the teleo-semantic tradition (Millikan, 1984). The goal of his framework is to provide a reductive account of the notion of representation that does explanatory work in cognitive science. The idea is that one can represent without representing *as such*. The former is a *purely referential* notion of representation. The latter requires either concepts or, for Burge, perceptual constancy.

One of the central concepts invoked in Shea's theory, and the one that best explains the olfactory phenomena discussed above, is that of a task function. According to Shea, a task function is a behavior F of some system S (organism or subpersonal mechanism) such that (1) F is a robust outcome function of S and (2) F is a stabilized function of S. These two components will be fleshed out below.

According to Shea, "Robust outcome functions are roughly those outcomes that result from behaviour which we humans are inclined to perceive as being goal-directed." (p. 52.) There are two components of robust outcomes. Specifically, "(i) the same distal outcome is produced in response to a variety of different inputs to the system; and (ii) the outcome is produced successfully across an array of relevant external conditions." (p.54.) As Shea notes, motor control of reaching is

paradigmatically robust (p.53); it adjusts to changes in the movement of the target object. More to the present point, though this will be explicitly defended in what follows, being able to detect odorants across a range of inputs is also a robust outcome function.

A quick note: Superficially, the notion of a robust outcome function bears significant similarity to the notion of perceptual constancy. Indeed, it might be remarked that robust outcome functions just are perceptual constancies. However, Burge claims that (i) "Not all selectivity with respect to registration of proximal stimulation constitutes the formation operation involved in a perceptual constancy. Some neglect or filtering of noise occurs in any well-functioning sensory system, perceptual or not. Non-perceptual sensory systems suppress some information in a signal to respond to information useful to the organism." (2010, 410.) This suggests that figure-ground segregation is not a genuine constancy, but as will be argued, it does fit squarely with the notion of a robust outcome function. And so, there seems to be good reason to distinguish them.

The second component of a task function is the concept of a stabilized function. This is roughly the explanation of the way a particular behavior or behavior type is stabilized over time. This stabilization results from evolution through natural selection, behavioral learning through feedback, contribution to persistence of an individual organism, and deliberate design.[13] (p.42.) Importantly, each process is said to lead to good consequences. We shall first say something about how evolved capacities, and then learned ones, lead to good consequences. According to Shea,

[13] For this paper, the stabilization of behavior that results from design will not be discussed.

[a]n output is generated because of the good effect it will produce when, in an organism's evolutionary history, outputs of that type have contributed to survival or reproduction. In that case the organism is producing this output now in part because of effects that the same type of output has had in the (evolutionary) past. (p.56)

This idea is firmly situated in the context of teleofunctional accounts of representation. However, his account of representation is more explanatory in that it also spells out the content of representation by appealing to the persistence of an organism, learning through feedback, and design. The persistence component is essentially the claim that an organism, qua *autonomous system*, self-maintains by generating conditions for its own persistence when it is out of equilibrium (p. 58.) No more will be said about this sort of stabilization because it is not obviously applicable to any of the putative cases of olfactory representation we have discussed or will discuss. Finally, Shea treats learning through feedback (reinforcement learning) as a stabilized function. We won't dwell on the details since, as Shea notes, there are many varieties of learning. But in broad strokes, the idea is that,

Like natural selection, reinforcement can lead an organism to produce O more robustly by better detecting the circumstances in which its behaviour is likely to produce O; by adopting new means for producing O in new circumstances; or by increasing the robustness with which it can produce a particular means to producing O. (p. 61)

Here, O is just some outcome. So successfully grasping an object in unusual circumstances counts as an outcome that results from learning.

Let us say a bit about how these stabilized functions might work. It is plausible that reaching behavior can be stabilized as the result of evolution or learning. For illustration, consider a case of learning. Successfully grasping an object while wearing prismatic glasses is almost impossible at first; however, one can learn

to accommodate the shift in visual perspective and thereby become successful[14]. Importantly, learning to grasp given changes in the orientation of the object or one's hand leads to good consequences. If the incentives are set appropriately, one can learn new skills. Moreover, smelling behavior can likewise be stabilized by learning, as we will see.

3.1 Olfactory Task Functions

Regarding some of the olfactory phenomena mentioned above, it seems as if there is at least one possible task function of olfaction on offer, namely, figure-ground segregation: an agent's ability to discriminate odorants (without necessarily using concepts of course) given a noisy (smelly) background. The alleged intensity constancy is also partly explicable in the minimalist framework. Although there has been much debate about the reality of intensity invariances in olfaction, this paper assumes that there is no such phenomenon since rejecting intensity constancies provides the strongest case against representationalism about olfaction[15]. The latter portion of this paper explicates how we might get something like an intensity constancy given certain cross-modal interactions or the application of concepts. However, it should be noted, that even if olfaction doesn't issue in intensity *constancies*, it plausibly represents odor intensity simpliciter.

The task function of olfactory figure-ground segregation is the successful discrimination of some odorant given a background of other odorants. Furthermore, the task function is robust in that the same odorant can be identified through a range

[14] See Shea (2018, p.53) for a detailed account of this process.
[15] See Barwich, 2019, 2020; Hahn, 1994; Mainland, 2006; Millar, 2019, Teghtsoonian and Teghtsoonian, 1982.

of inputs. That is, given various signal to noise ratios, an organism can discriminate odorants from a background of odorants. Moreover, there seem to be at least two ways in which olfactory figure-ground segregation is stabilized; (1) through natural selection and (2) through learning. There is evidence showing that both are real.

Consider natural selection first. A rat can discriminate the odorants produced by a cat in a wide range of contexts, and if it could not do so, it would likely not survive. Moreover, it has also been found that rat pups reared independently of their parents have a fear response to the odorants produced by a cat. (Panskepp, 1998, p.19.) The idea here is that rats have an innate capacity to discriminate odorants produced by a cat without learning. The likely conclusion is that natural selection endowed rats with an innate capacity to discriminate cats (or other such predators) from other animals.[16] More precisely, the fear response that issues from cases such as these is likely what drove the stabilization of their ability to discriminate cat smells. Importantly, if they are to be successful (avoid being eaten), they must be able to distinguish a cat (again, the odorants produced by the cat) from a background of odorants across a range of contexts. For instance, rats must be able to smell odorants produced by a cat given variations in the concentration of whatever odorants are relevant to identify cats or given variations in the background in which the cat odorants are embedded. And presumably, having robust fear responses to such information likely reinforces their capacity to discriminate the smell. This is what makes the capacity robust.

[16] To be clear, the claim is *not* that rats utilize the concept CATSMELL, but rather rats represent the odorant produced by a cat purely referentially.

In contrast, consider the case of expert perfumers. For expert perfumers, the ability to discriminate certain simple odorants given a complex structure (e.g., a perfume) primarily results from learning. The suggestion is not that *nothing* about this skill is innate, however. Rather, the claim is that being a successful perfumer requires proper education and training in smelling perfumes (discriminating a perfume's component odorants). Indeed, there is quite a bit of evidence that learning affords people with finer-grained discriminatory abilities, when it comes to detecting odor information in an odor complex, than their untrained counterparts. (McGann, 2015; Fournel et al., 2017; Mourquecho-Campos, 2019.) For example, according to a study by Mourquecho-Campos (2019), after eight training sessions, subjects were better able to disassociate binary odorant mixtures than the non-trained group. These abilities must be robust to explain their general skill of identifying underlying odorants. Although employing concepts are plausibly relevant in these cases, they need not be more generally. There is a plausible case to be made for the claim that affective learning would be sufficient to drive one's discriminatory capacities in a great deal of cases. For instance, vomiting when presented with a smell leads to the negative appraisal of the stimulus, which makes it more likely that the smell will be noticed in the future so that the creature can avoid it quickly to prevent vomiting. This stabilized response to certain smells can therefore underwrite one's capacity to discriminate it from a background.

The rat example and the expert perfumer example make a plausible case for the idea that olfactory figure-ground segregation is robust in the relevant sense, as well as stabilized in the relevant sense. Both behaviors are robust according to Shea's

framework insofar as the capacities are flexible. As we saw, rat pups can presumably identify the scent of a cat across variations in the ratio of the target odorant to a background set of odorants. If they could not do so it is unclear that rat pups could be successful in environments in which the ratio of cat smell to other smells varies considerably. Moreover, since this capacity is innate, the best explanation is that it is stabilized by evolutionary pressures. The expert perfumer's capacity to identify certain odorants is also presumably robust in a similar way. Likely, people are trained to identify odorants in fairly narrow contexts, but if one is to be a successful perfumer, one must be able to identify the same components when they are embedded in various mixtures of various ratios. Likewise, their capacity is presumably learned (and possibly partly innate), so it also satisfies Shea's requirement for being a task function.

Implicit in one's ability to discriminate an odor from a background is one's ability to represent the intensity of the target odorant. For instance, one can discriminate an odor from a background given differing ratios of the target odorant to a background. And presumably the various representations issue in distinct experiences. Now, this does not yet amount to the representation of an *invariant* intensity. That is, representing an odorant's intensity does not yet mean that olfaction is representing a constant and invariant feature of the world. For instance, the apparent intensity of an odor, plausibly understood as a representation of the total amount or concentration of odorants, is consistent with either part of the world being intense to degree N or simply one's experience being vivid to degree N. The point is simply that olfaction represents the concentration of an odorant relative to a

background. It represents the degree to which a target smell differs from the background. Section 4, however, shows how representing some concentration of odorants can, when embedded in an object-file, issue in an invariant representation of intensity.

As we just saw, Shea's account of a task function is apt to explain some olfactory capacities. This puts one in a good position to see that olfaction is minimally representational, in the sense employed in cognitive-science explanations, without involving on perceptual constancies. Moreover, once these minimal representations become components in a more robust cognitive process, we can see how we can get olfactory objectual representations.

3.2 The Virtues of Minimalism

Before moving on to a constructive account of olfactory objectual representation, it is important to consider an objection. One might complain that the view set out in this paper crucially changes the subject away from Burge's concerns. Burge, and those who presuppose a rather strong account of representation, might complain that this change of subject really fails to get at what matters. They might suggest, for example, that perhaps there are minimal olfactory representations, but that isn't the sense of representation that matters for psychology. Essentially, the objection boils down to the claim that we can quibble about what "represent" means, but, really, objectual representation is the relevant category: it is the one operative in psychology. There are several responses to this worry.

First, Robert Rupert's (2018) overview of much of the recent debate about representation suggests that there are many theories of representation on offer, yet

none seem to be as stringent as Burge's. That much of the literature on representation does not require that representation disclose to us a mind-independent objective reality suggests that there are good independent grounds to be skeptical of the notion of representation presupposed in the olfaction literature. The diversity of theories of representation thus provides at least some reason not to be worried about the present changing of concerns.

Relatedly, as Barwich (2019) notes, the notion of objectual representation (though she doesn't use that term), the one at issue in these debates, is derived from vision science (p.1, 2019.) So even if that notion of representation is consistent with vision-based psychology, it is overzealous to impose the same constraints on another successful, albeit nascent, subfield. In fact, since minimal representation can be shown to explain many olfactory phenomena, we seem to have an alternative, and perhaps more general, psychological theory of representation.

Third, Shea notes that objectual representations (though he also doesn't use that term) are apt to explain many personal level representations, such as one's beliefs and desires, but many of the phenomena discussed in the literature are simply about non-belief-like (subdoxastic) representations. (See Li et al., 2007) Indeed, many of the examples set out by Shea (2018) have this character also. But moreover, many are also genuinely psychological. (Shea, 2018, pp. 174–175).

Finally, informational accounts of representation are ubiquitous in scientifically informed discussions of the mind.[17] So, although appealing to minimal

[17] For instance, see Aydede & Güzeldere 2005; Dretske, 1981; and Fodor 1987 for informational accounts of representation.

representations is, in an important sense, a change of concern, it is not unorthodox and there is much support in the literature.

All these considerations suggest that we should be ecumenical about the concept of representation and allow that there are different senses of the term that are meant to explain different things. The important takeaway is that objectual representation is too strong a notion to be a general account of representation. It is certainly apt to explain *some* psychological phenomena, but not all. Indeed, since olfactory representations are often invoked in the psychological literature, rather than dismissing them, the goal should be to provide an alternative account of representation that fits the practice. The foregoing discussion is meant as a first step towards that goal. However, the remaining discussion will be an attempt to vindicate the idea that there are, at least at the personal level, some olfactory representations with olfactory objectual properties as content. Importantly, however, representations of that sort are explained in part by mechanisms outside of olfaction itself.

Now, there is an additional yet related worry.[18] Perhaps it is true that the notion of minimal representation ought not to be dismissed for the reason just rehearsed, but one might remark that this doesn't yet show that minimal representation is the notion that matters for those who study *perception*. The concern is that the study of perception, as opposed to mere representation, really does require that one account for the *objectual* aspects of representation. Another way of putting it is that maybe perception does in fact depend on representing an objective world. Although this worry is plausible, it does presuppose Burge's notion of perception. It

[18] Many thanks to an anonymous reviewer for pushing this point.

is true that some perceptual modalities represent an objective world all on their own, but that doesn't show that *all* perception must be like that. That consideration might simply lead one to think that there is perception, on the one hand, and something else, more minimal, on the other. This very well might be true, but it does seem to simply reassert the assumption of external directedness that many have worried about. So, even though olfaction and some other modalities (vision, audition, and touch) are different in crucial ways, it doesn't immediately follow that some are examples of perception whereas olfaction is not.[19]

Section 4 : Object-files and objectual representation

It seems that some of the phenomena meant to vindicate the representation of olfactory objectual properties are explainable using the framework of minimal representation. This means that although olfaction does not issue in perceptual constancies, and thus does not issue in olfactory objectual representations, olfaction is nevertheless representational in a minimal sense. This conclusion poses a puzzle, however.

Since we intuitively do in fact represent olfactory properties as objectual, as when we represent a smell as the smell of a rose or as the smell of a perfume, it seems we need an account of olfactory representations of that sort.[20] A minimal account of

[19] Alternatively, one might think that part of what distinguishes perception from imagination is that perception necessarily involves the awareness or seeming awareness of worldly things in one's environment, whereas imagination does not. However, it is also plausible that the distinction between imagination and perception might instead be partly grounded in the feeling of control. This distinction is orthogonal to the distinction between representing an objective world or not. For instance, one might imagine a situation that seems to be as of the objective world, yet one might distinguish it from a perception if the content seems to be in their control (e.g., as when one looks for familiar forms in clouds).

[20] Now, one might remark that there is no intuitive objectual representation of smell if we do not have any cross-modal or verbal cues to an object. I agree with this claim; however, my reason for endorsing

representation thus seems inadequate for a general account of olfactory objectual representation. The task of this section will be to explain how this might work by appealing to object-files, feature-binding, and attention.[21] The suggestion is that, because of a post-perceptual process, creatures can represent olfactory objectual properties. Consequently, we will see how this can be used to explain olfactory filling-in and the representation of an invariant intensity.

The first phenomenon of concern is just the idea that odors can be objectual; it is the idea that odors seem to be in the world, at least sometimes.[22] The second is olfactory filling-in and the third is the alleged intensity constancy. It will be argued that olfactory filling-in and intensity constancy partially depend on representing objectual properties in the first place, and so we will begin with an explanation of the latter. To provide an adequate explanation, it first needs to be shown how this might work in other, well-established domains. The idea to be defended is that in general, perception builds object-files (objectual representations), and this depends on

---

the intuition is that there are almost always cross-modal or verbal cues to objects. I want to thank an anonymous reviewer for pushing me on this point.

[21] The concept of an object-file is very similar to the concept of perceptual categorization presented in Barwich (2020) and the profiles account of Smith (2017) since they all account for cross-modal interactions. However, object-files can also depend on high-level cognitive processes that seem to be absent in the case of Smith. For instance, in a review of Keller (2016) he notes that Keller's theory, "invokes an unnecessary high-level cognitive process to solve a problem that is often the result of crossmodal associations between inputs from different sensory systems and experience of how they typically combine in everyday episodes of multisensory perception" (p.794.) Below I argue that something like an unconscious inference is in fact necessary to bind modality distinct properties together. Importantly, Barwich (2020) invokes a very similar account. However, her view emphasizes that olfactory perception is a skill, whereas I emphasize that olfaction is representational through and through. Moreover, Barwich argues that this process does not explain the formation of olfactory objects. I argue, on the other hand, that in virtue of the objectual nature of the properties that are bound up with odor representation, we do in fact get the representation of olfactory objects. So, although perceptual categorization and object-files are alike in spirit, they do have some small but important differences.

[22] It should be noted that many philosophers do not share this intuition. In fact, many argue that smells are mere subjective feels (see Burge 2010). Since the goal of this section is to vindicate the intuitions of some, namely those who do think that smells can be objectual, I take this section to be a defense of their position.

attention and feature-binding. Moreover, since perception is widely agreed to be constructive, we should expect the same results to hold in the case of olfaction.[23]

As we saw, olfaction itself does not issue in objectual representations, that is, representations of olfactory objectual properties. However, there do *seem* to be olfactory contents of that sort. At the very least, phenomenologically, it is natural to think of objects of our awareness *seeming* to have some smell or other. Since properties of this sort do not arise because of olfaction itself, there must be a cognitive process that somehow constructs representations of this sort.

The claim to be defended is that olfactory objectual representations are constructed when olfactory minimal representational content becomes bound to contents contained in object-files that *do* meet the criteria for genuine objectual representation (those derived from perceptual constancies). In these cases, an olfactory objectual property is parasitic on the properties that derive from other senses. Crucially, this constructive process depends on cognitive processes which are general and independent of olfaction itself, in particular, attention, inference, and, sometimes, the imposition of concepts. To defend this position, we need to consider Anne Treisman's *Feature Integration Theory* (FIT). Furthermore, it is important to review some interesting and recent applications of her theory to modalities other than vision to see how this might work for olfaction.

Treisman's theory is a multi-stage theory of perception. At the first stage, information about distinct features (color, shape, texture, etc.) are represented

[23] Although I know of no data directly in support of olfactory object construction, the argument has the form of an inductive inference. Importantly, visual/auditory forms of object construction are well established, and thus provide a firm foundation on which to base my argument.

separately. This stage happens extremely quickly and automatically. At the next stage in the process, a subject combines these features into a unified representation of an object, and in some cases, into a unified representation of an event[24]. This process is known as feature binding. Visual examples abound, but there are clear cases of cross-modal feature binding as when one notices that a particular bird song is the song of a visually represented bird. The song and the visual components connect in a way that makes them seem as though they are about the same thing, namely, the bird. This process requires attention to the features of interest. Although the exact mechanism is still controversial to some extent, most everyone accepts that attention is "...required to glue, or bind, visual features together effectively at particular locations within a putative mastermap of locations." (Spence & Frings, 2019, p.364) There also likely needs to be other processes at play during a token binding event such as the deployment of an inference about the likely relationship between the stimuli. A toy principle of this sort will be suggested below.

The rubber hand illusion is perhaps the most well-known case of cross-modal binding. "Watching a rubber hand being stroked, while one's own unseen hand is synchronously stroked, may cause the rubber hand to be attributed to one's own body, to 'feel like it's my hand'." (Tsakiris & Haggard, 2005, p.80.) Crucially, the visual information and tactile information need to be attended to within an appropriate temporal window for the effect to occur. Outside of that range, the effect disappears. Of course, the rubber hand illusion is a case in which the feature-binding is non-

[24] In contexts in which the low-level information is constitutive of a temporally extended sequence (say, a melody), then the information is encoded into a single event file, and not an object file.

veridical, but this also happens when you watch someone stroke your real hand. In the veridical case, the movement and the touch are both integrated into the same event.

There is some experimental evidence of cross-modal interactions in the case of olfaction as well.[25] For instance, Gottfried & Dolan (2003) showed that "olfactory detection was faster when odors appeared in the context of semantically congruent visual cues" (p. 380.) However, the authors note that the visual cues on their own were not sufficient to explain the effect. Rather, the effect depends on access to semantic information that is consonant with the olfactory stimulus. For example, the effect occurred in cases in which pictures of ice cream or cake were presented alongside the *vanillin* odor, and not if the image was semantically unrelated to the odor. This finding might provide some reason to think that this is not genuine binding since it is unclear that semantic congruence is a necessary condition for binding in general. For instance, in mundane case of visual feature binding color is bound to shape, but colors are not the sort of thing that could agree *semantically* with shapes. However, that research shows that there are at least cross-modal effects in the case of olfaction.

In any case, the following captures how binding might work in the case of olfaction. Odorants are first minimally represented in the way suggested above, and content contained in those representations becomes bound into object-files that have as content certain properties represented by objectual representations, namely properties which result from perceptual constancies. This happens as the result of attention being simultaneously directed at smell contents contained in minimal

[25] Many thanks to an anonymous reviewer for bringing this work to my attention.

representations as well as to other properties that result from perceptual constancies. It is by making use of an inference, about the likely source of the olfactory content, that attention is also directed to the source, whatever it turns out to be (perhaps some visual feature). Moreover, this can happen very swiftly and unconsciously as when someone shoves something smelly in your face. Intuitively, this account can make sense of the various cases in which it seems as though we are representing olfactory properties as features of the objective world, namely, cases where we explore our environments to glean new information about the source of smell information. Perhaps more commonly, however, it also allows us to explain cases that begin with the visual perception of an object, say, but whereby one consequently sniffs it to determine its smell.

Now, consider an example, namely, the unpleasant experience of walking into a store such as Yankee Candle. Odorants are extremely salient and are also quite overwhelming. In this case, although many different odorants are present, they do not seem localized, i.e., they are diffuse and so do not seem like objectual properties in the relevant sense[26]. However, when one explores the store and takes isolated sniffs at particular locations, the smells then seem to be properties of *particular* objects. Crucially, the deployment of a conscious or unconscious inference seems like a natural way to explain how one comes to represent the *source* of an odorant.[27] For example, one might use a principle which says that 'if the intensity of smell S changes as a function of the proximity to a visually represented token object O, then S

[26] They no doubt capture our attention, but as we saw in our discussion of figure-ground segregation, attentional capture is insufficient for object perception.
[27] Barwich (2020), drawing on Hermann von Helmholtz also argues that unconscious inferences are crucial to explain olfactory experiences. For instance, see Chaps. 8 and 9.

is part of O'.[28] Before exploration and the use of the principle, smell is minimally represented. But once that content is bound to properties, which are objects in Burge's sense, a smell (i.e., odorant) is objectually represented. These cases appear to be quite common and seem fairly simple to explain, but there are other more complicated cases. Specifically, the "filling-in" and intensity constancy phenomena expounded by Millar (2017) are such cases. Both phenomena will be explained in turn.

Recall that when people are presented with a corrupted odor compound—a compound that is missing specific odorants—they can still perceive its *type*, say, COFFEE. This case is related to the above Yankee Candle case but is distinct in that it seems to require concepts. Imagine the experience of smelling perfectly roasted coffee. The smell is quite distinct from other smells and is distinguishable from the simultaneously presented smell of, say, bacon. The suggestion is that the coffee smell is only objectually represented when one has explored one's environment and has consequently bound the smell of coffee to the objectual representation of the coffee that does exploit perceptual constancies (e.g., the visual properties coffee).[29] Now imagine the same scene but where the coffee is over roasted such that it is burnt (the odorants are corrupted in the way set out by Millar). In this case, the capacity to

[28] This is not supposed to be an exhaustive principle. For instance, there likely needs to be a "no-association" principle employed as well to eliminate mere correlative effects. For instance, perhaps the smell of black pepper is always accompanied by the visual representation of salt, but it would be a mistake to bind the smell of pepper to the visual representation of salt. Since ordinary cases of olfactory binding are veridical, these purely associative cases need to be excluded somehow. Perhaps binding requires a counterfactual condition as well to rule out these problem cases. This article makes no commitment to how this should work, but it should be noted that this is a general problem with feature binding and not one exclusive to olfaction.

[29] The claim is not that vision is necessary, but rather that any sense modality that issues in representations in Burge's sense is necessary.

distinguish it from the background is still present (perhaps more so than in the non-burnt condition). However, what makes the representation as of coffee? That is, why is the corrupted compound represented *as coffee*? The problem is that the burnt condition is so different from the normal condition that there seems to be an explanatory burden on the person who argues that olfactory filling-in is purely perceptual. It seems that nothing about the smell itself or olfaction itself can explain it. In fact, it seems that only when you have the concept COFFEE, can you objectually represent both, quite distinct, smells as being an instance of the same type. The thought is that the concept COFFEE is an object-file that contains as content the distinct and likely *disjunctive* ways that the thing in question might smell. And it should be emphasized just how disjunctive the smells could be. In other words, there is a huge range of inputs that might cause the tokening of the objectual representation of COFFEE, where each input might bear *no* similarity to any of the others.[30] In order to successfully represent all the various ways coffee could seem, one would need to embed minimal representational content into the coffee object-file, thus creating an olfactory objectual representation of coffee. In virtue of this embedding process, one can still represent the corrupted coffee (burnt coffee) *as* coffee.

Furthermore, representing that an odorant's intensity as invariant might work similarly. For example, once minimal olfactory content is embedded in a COFFEE object-file, one might represent that the coffee's intensity hasn't changed even if the subjective intensity has. Perhaps one might think to oneself that coffee is not the sort

[30] Besides smelling burnt, coffees can also be fruity, flowery, nutty, spicy, herby, caramelly, chocolaty, carbony, resinous, or any combination. Moreover, those are just the superordinate categories and each one is composed of several subcategories.

of thing whose properties can change at the rate at which adaptation sets in. The idea is that the category COFFEE, and basically any other categories for things in the world, are such that they pick out stable properties. The upshot is that even if one's subjective experience of an odorant's intensity varies, one might still experience the worldly entity as unchanging. However, if true, this would clearly be the result of a cognitive process that extends beyond the proper functioning of one's olfactory system. Nevertheless, such a process would still issue in a stable representation of an odor's intensity.

We have just seen how olfactory minimal representational content might combine with other sensory content such that the resultant representation will objectually represent olfactory objectual properties. The suggestion is grounded in standard cases of cross modal feature-binding and the application of an inference. Moreover, it also seems equipped to explain both olfactory filling-in and apparent intensity constancy in a way that does not presume that there are olfactory perceptual constancies.

Before concluding, we should first note a plausible worry with the present account of olfactory perception.[31] One might think that the above treatment of olfactory construction depends too heavily on either experiencing the source of a smell from another modality or on employing a concept. Perhaps representing olfactory objectual properties occurs much more automatically. More precisely, one might think that searching for the source of an odor presupposes that there is some object or other producing it in the first place. For instance, suppose one enters their

[31] I thank an anonymous reviewer for pointing this out to me.

kitchen and amongst all the smells, one singles out a particular novel foul odor. Without employing a concept, beyond the demonstrative "that smell", or seeing the source of the smell, one might think that there must be some objectual property.

There are two responses to this concern. First, as we saw above, figure-ground segregation does not amount to the representation of objectual properties. Just because there is some information that captures one's attention, it doesn't follow that one is therefore representing an objectual property. Second, intuitively, having a smell experience does not, by itself, present anything objectual. That is, a smell experience, prior to following a gradient of olfactory information, does not seem to be about anything in the world. Often such smell experience can strike one as being perfectly hallucinatory in nature. The contention of this section of the paper is that we get a transition from a purely subjective olfactory quality to the representation of objectual properties only when one either employs a concept or searches for the source of the experience. Now of course, one might still worry that unless one assumed that the reason one is having a smell experience is *because* there is some object or other causing it, one would not search. However, it is one thing to *believe* that the cause of one's experience must be outside of oneself, and quite another for one's perceptual experience to be objectual. It seems that the debate that has concerned many has to do with the status of *perception* in the first place. Simply showing that one ordinarily has certain beliefs about the world and one's causal relation to it doesn't undermine the present proposal.[32]

[32] In a similar vein, one might insist, however, that it is just as plausible that olfactory phenomenology is intrinsically objectual. For instance, Richardson (2013) argues that the role of sniffing can explain the objectual (though she doesn't use this term) nature of olfaction. She claims that "the involvement of sniffing in smelling is manifest to us in the conscious character of olfactory experience. Olfactory

5. Conclusion

This article has argued that olfaction does not issue in objectual representations. This is because there are no genuine olfactory perceptual constancies. At most, olfaction represents in a *minimal* way in virtue of extracting signal from noise or perhaps by representing odor concentration. However, there are also cases in which olfactory phenomenology seems to be about olfactory objectual properties. In these cases, one objectually represents smells as properties of the world because of cross-modal feature-binding and/or the employment of concepts. Given that olfactory representations sometimes do have this more robust character, an account of the exploratory and constructive nature of olfactory perception was surveyed. The moral of the story is that olfaction does not represent objectual properties, but *we* do.

---

experience does not just involve awareness of certain qualities, such as vanillaryness or mustiness or whatever. It also involves odours that have these qualities seeming to be brought into the nose when we sniff." (p.11) The claim seems to be that sniffing behavior underwrites the externally directed phenomenology of olfaction. So, although the account on offer in this paper is plausible, the alternative, based in sniffing behavior, has not been ruled out. Now, one might respond to this by simply noting that we can sometimes have smell experiences that are not the result of sniffing. An example would be if someone shoved something smelly under your nose. The claim is not that there can be smelling without breathing, but rather, that there can be smelling without sniffing. Sniffing then, might very well explain the objectual nature of some olfactory experiences, but it certainly is not a necessary precondition of all olfactory experience. If true, then sniffing behavior does not allow one to conclude that olfactory experience is itself objectual. Importantly, however, the role of sniffing behavior in explaining the externally directed nature of olfactory experience is consistent with my positive account of olfactory experience. The main difference is that I argue that exploratory behavior in general, including sniffing (see the Yankee Candle example above) is not strictly speaking a constituent of olfaction. Thanks to an anonymous reviewer for pushing me on this point.

Chapter 3:
Additional Resources for Sparse Theories of Phenomenal Consciousness

1. Introduction

Over the past several years there has been an active debate about the richness of experience; the so-called phenomenal-overflow debate. The debate centers around the question of whether or not the bandwidth of phenomenal consciousness is greater than the bandwidth of access consciousness. Those who think phenomenal consciousness is richer than access consciousness defend the "rich view", the main spokesperson of which is Block (2011, 2014).[33] Those who think the contents of phenomenal consciousness are as sparse as access consciousness defend the "sparse view". The main proponents of this view are Cohen and Dennett (2011), Cohen, Dennett, & Kanwisher (2016), and Dehaene (2014).[34] By "phenomenal consciousness", Block and others are simply referring to the subjective, felt quality of experience—the what-its-like quality of experience; whereas "access consciousness" is taken to be the availability of the contents in question to systems for reflection, action planning, verbal reporting, and so on—those subsystems that are often mentioned in the global broadcasting literature.

[33] See also, Aru & Bachman (2013); Bronfman et al. (2014); Koch & Tsuchiya (2007); Lamme (2010); and Usher et al. (2018) for additional contributions to the rich view.
[34] See also, Fazekas & Nemeth (2018); Naccache (2018); Odegaard et al. (2018); Overgaard (2018); Sergent (2018); and Ward (2018) for additional contributions to the sparse view.

Block argues that phenomenal consciousness is associated with information encoded in a fragile short-term sensory-memory store, which has a much larger capacity than working memory. The difference in encoding capacity partially grounds the view that phenomenal consciousness overflows cognition. The thought is that one is phenomenally conscious of more than one can report because contents encoded in sensory memory are much greater in extent than those available to access consciousness. Cohen, Dennett and Kanwisher (2016), on the other hand, argue that one is phenomenally conscious of just whatever items are represented in working memory, and that any additional impression of richness[35] results from consciously represented background statistical information, as well as from continual and unconscious eye saccades and shifts of covert attention. Much like whenever one opens the refrigerator door and the light is on, this idea suggests one has the impression of richness and vivid detail because whatever one attends to is phenomenally conscious.

Recently, both proponents and opponents of phenomenal overflow have utilized scene statistics in an attempt to ground their particular view. Scene statistics, also known as summary statistics, gist representations, or ensemble representations,[36] are mental representations that have high-level abstract properties as their contents. For instance, scene-statistical representations might contain as content the diversity of coloured items, that there are many stripes (but without specifying a determinate number of stripes), and so on. Since scene statistics have been appealed to both in

[35] There is a sense in which there are two explananda in the debate; the impression of richness, on the one hand, and the reportability of the impression of richness on the other. However, the evidence for the former is often put in terms of the latter, and so it seems it is best to focus on reportability.
[36] These terms will be used interchangeably; nothing rests on the distinction between them.

arguing for and against phenomenal overflow, this naturally poses a problem for the debate because it remains unclear what scene statistics are actually equipped to explain. There is at least one important question about scene statistics that might sway the debate one way or another: are scene-statistical representations rich or sparse? In what follows, it will be shown what bearing an answer to this question has on the debate.

Following Cohen and Dennett (2011) and Cohen, Dennett and Kanwisher (2016), this paper will motivate a sparse interpretation phenomenal overflow. As a result, it will be argued that the application of scene statistics is often apt as a response to various phenomena discussed in the literature (namely, Sperling-type experiments) — where some items are perceived with rich detail and the remainder are represented by a summary statistic. However, Cohen, Dennett and Kanwisher (2016) explain the subjective impression of richness by appealing not only to scene statistics, but also to seemingly effortless eye saccades and the consequent 'refrigerator-light illusion'. In what follows, it will be argued that this is correct for the most part, but there are other tools a sparse theorist can invoke to make their explanations more compelling. First, it will be argued that feature binding can explain additional phenomenal riches. That is, feature binding can explain how one might become phenomenally conscious of more information than is typically thought. Second, it will be argued that the refrigerator-light illusion is perhaps not a singular phenomenon, but rather might occur in several ways. As a result, an alternative way in which effortless shifting of attention can occur will be explored.

This paper will proceed as follows. First, the overflow debate, along with the key studies invoked by both proponents and opponents of phenomenal overflow, will be introduced. Second, a minimalist conception of scene statistics will be set out and motivated. Third, several additional sources of richness and *apparent* richness, that could be utilized by an opponent of overflow, will be described. The tentative suggestion is that they succeed. However, at a minimum, they are factors everyone in the debate should consider.

It should be noted at the outset that both components of the view set out in this paper are consistent with the sparse view put forward by Cohen, Dennett and Kanwisher (2016). The sparse view, generally understood, merely suggests that the bandwidth of phenomenal consciousness is identical to the bandwidth of access consciousness. It does not entail anything about what the actual capacity limit must be. Of course, the common view is that the bandwidth is 3–4 items, but the details of this claim are thoroughgoingly empirical and there does seem to be a lot of evidence that the 3–4 items can each contain quite a bit of information, as will be shown. The view that will be pushed is therefore a kind of sparse view in that it suggests the capacity limits of phenomenal consciousness and access conscious are identical, but the amount of information contained in working memory can be higher than is assumed. But perhaps less controversially, the second proposal is simply that there is more than one refrigerator-light illusion. Empirical evidence for these claims will be introduced.

2. The Phenomenal Overflow Debate

The overflow debate is concerned with common-sense intuitions about the apparent richness of experience, as well as with people's discriminatory successes within several experimental paradigms (i.e. being able to correctly identify the constituents of various arrays and scenes). The main focus of this paper will be on the latter, though much of what will be said will apply to the former as well. Everyone in the debate agrees there is the impression of richness. As a result, there are several ways in which one might handle that impression. Either one tries to explain it, as does Block, or one tries to explain it away, as do Cohen, Dennett and Kanwisher (2016). In what follows, a hybrid view, suggesting there is more richness than is assumed by the sparse theorists, but less than by rich theorists, will be explained.

2.1. Block's Claim of Richness

The most prominent case to be made on behalf of the rich view comes from George Sperling's paper 'The Information Available in Brief Visual Presentations' (1960). In the experiment, subjects were presented with an array of letters arranged in a 3 x 4 grid for a brief period (anywhere from 15 ms to 500 ms) and were asked to recall the content. When subjects were asked to recall every letter, they could only report 3–4 (roughly the capacity of working memory). However, when any given row was cued by an auditory signal after the image was offset (high, medium, or low tone, corresponding to the high, medium, or low row of the array, respectively), subjects could report about 3.5 letters per row, adding up to 10.5 letters. Block takes this as evidence in favour of overflow because it shows that the content from the entire grid was still available for report. Block and Sperling call this the 'partial report

advantage'. Specifically, Block argues that each letter is consciously represented in virtue of being encoded in fragile visual short-term memory (FVSM). FVSM is a type of sensory short-term memory which consists of persisting visual representations lasting around 4–5 seconds (Block, 2011). According to this view, each letter in the array is consciously represented in FVSM, which is what explains the subjective impression of seeing all the letters.

Another study invoked by Block is Bronfman *et al.* (2014). In this series of experiments, Bronfman and colleagues utilized a variant of the Sperling paradigm. Instead of just letters, subjects were presented with a 4 x 6 array of coloured letters where one of the rows was precued. After the image was offset, there was a brief blank display, and a cue directed the subjects' attention, not to a particular row, but to a particular position in the row. Furthermore, and crucially, in some trials, subjects were also asked to express the colour diversity (high or low) of a specific cued or un-cued row. It turned out participants could register colour diversity in the un-cued rows, which suggests colour diversity becomes available to a subject 'cost free', attentionally speaking. Bronfman *et al.* suggest that colour diversity is encoded as a consciously accessible summary statistic.

Block and Bronfman take this as evidence in favour of phenomenal overflow because they think the statistical representation (high/low diversity) cannot be determined unless the colour of each letter is individually and determinately represented, and that these representations are also conscious. More will be said about this interpretation in what follows.

To summarize: the main point to consider about Block's interpretation of Bronfman's study is the assumption that '[e]nsemble perception of the kind reported by Bronfman reveals that there must have been conscious awareness of specific colors beyond the limits of the global workspace because a trace of that conscious awareness in the form of a diversity judgement can enter the global workspace for free' (Block, 2014). Block's and Bronfman *et al.*'s claim is that the diversity judgment could not be available to a subject unless each of the items that were computed over (to create the statistical representation) were conscious at some point. This point is important because Cohen, Dennett and Kanwisher (2016) also utilize statistical representations, but they do not think the generation of the statistical representation requires that the individually represented colours were conscious when computed over to form the statistical representation. Crucially, for Cohen *et al*. the statistic is conscious, but the content that was 'averaged' to get the statistic is not. In what follows evidence for the latter claim will be provided. However, it will be argued that a sparse view has further strengths, and can accommodate higher degrees of richness, than is generally allowed.

3. The Sparse View — the Details

3.1. Scene Statistics

Part of Cohen Dennett and Kanwisher's (2016) scene-statistical explanation for the Sperling data is that most of the individual letters are *unconsciously* represented. The unconscious representations can be encoded in FVSM, but those contents are not made conscious unless attention is directed upon them. Their idea is

that the cue directs attention towards those unconsciously represented letters and, as a result, those representations are made conscious. This explains the ability of the subjects to accurately recall some of the letters in the 'partial report advantage' condition. However, the lingering impression of seeing all the letters is explained away by positing a generic 'letterness' representation. In other words, the particular distinguishing features of each letter — whether there is an 'A' or 'S' — are lost (not consciously represented), and what remains is an abstract (high-level) gist representation. The representation can be formed by simply finding a consistent abstract property among all of the unconscious representations, or just among some region of the visual field, namely that they are letters, and merely consciously representing *that* property, in addition to a small number of particulars.

As a result of Mack and Rock's (1998) seminal study on inattentional blindness, it was thought that scene-statistical representations were consciously available to a subject even if their attentional resources were directed elsewhere; say, to some items directly in front of them. Consequently, the commonly held view was that scene-statistical information was available for reporting independently of attention; the statistical information was thought to be available 'cost free'. However, evidence gathered by Cohen, Alvarez and Nakayama (2011) called this assumption into question.[37] Specifically, they found that the initial studies were flawed because

[37] For additional studies showing non-focal items are not available 'cost free', see Slagter *et al.* (2007), Stein *et al.* (2009), Mack and Clarke (2012), and Jackson-Nielsen, Cohen and Pitts (2017). Jackson-Nielsen, Cohen and Pitts (2017), in particular, showed that, *contra* Bronfman *et al.*'s (2014) findings, colour-diversity representations are not available cost free. This is important for the overflow debate because it provides more evidence that attention is required to perceive summary statistics. I thank an anonymous reviewer for bringing this study to my attention. One thing to note, however, is that it does not engage directly with one of the main conclusions of Bronfman *et al.*'s study which is that the *components* of the summary statistics are phenomenally conscious *if* the summary statistic is conscious. Thus, there are two orthogonal dimensions to the study: one which says conscious summary

they did not sufficiently strain attention. In other words, the original studies failed to tie up all of the subjects' attentional resources. Once attention is properly controlled for, it seems attention is, in fact, crucial to reporting on background information, i.e. noticing and making judgments about the content of the scene beyond focal attention. In the study, subjects were given various attention-demanding tasks, for example MOT (multiple object tracking) and RSVP (rapid serial visual presentation) tasks, and were asked to indicate if they noticed the background shift. The subjects were found to be inattentionally blind to natural scene statistics. Because their attention was being utilized for a quite effortful task, they systematically failed to 'spread' their attention to the background. This finding provides us with *prima facie* reason to reject one of the main theses of Block and Bronfman *et al.* — that scene-statistical information is available cost free.

However, even if it is established that scene-statistical information is not available to a subject independently of their attentional resources, it still remains unclear whether the items that were computed to form the statistic need to be conscious.

So, what is the nature of scene-statistical representations? They are described by Cohen *et al.* as follows: 'the ensemble representations, or summary statistics, are formed by collapsing across the measurements of individual items to form a singular

---

statistics are themselves independent of attention, and the other which says the individual components, that are computed to generate the statistic, are conscious if the summary statistic is conscious. For this paper the focus is on the second dimension. The reason is, even granting that summary statistics do not require attention, it is a much stronger claim that phenomenally conscious summary statistics entail that the items that compose the statistic are likewise phenomenally conscious. It is this stronger claim that is of interest to me in this article.

description of the group. Although items that are not focally attended to are represented with poor resolution, averaging across these imprecise representations allows the system to obtain an accurate measure of the entire group' (Cohen, Dennett and Kanwisher, 2016, p. 327).

There are two important takeaways from this passage: (1) the individual items are not consciously represented with high resolution and (2) there is an accurate measure of the whole group, but not each item. Interestingly, they argue that, although people see more than the 3–4 item limit of working memory, they will see less than they think they do. In order to establish this, Cohen and colleagues accept the conclusions of contemporary models of attention and working memory, but argue that, since attentional resources are almost never completely utilized, there will often be some leftover attentional resources directed at representations that exploit redundancies in the environment. In their words:

> The world beyond focal attention is not in darkness because when attention is not entirely engaged by a primary task, and it is unclear if attention can ever be entirely engaged using psychophysical techniques, excess attentional resources are automatically deployed elsewhere... Focal attention often leads to high resolution percepts whereas the percepts from distributed attention are at a lower resolution but with certain basic elements preserved. (Cohen, Alvarez and Nakayama, 2011, p. 360)

So far, there are some aspects of scene statistics that should be agreed upon by most everyone. However, additional claims made by some seem to be entirely unmotivated. For instance, as mentioned above, Bronfman *et al.* (2014) claim that ensemble statistics actually support overflow because individual colours summarized in those statistics were first represented consciously. This is an odd claim to make considering the fact that unconscious colour priming effects are quite robust, and that even blindsight patients can in fact unconsciously represent colour information. For

instance, two such studies suggest colour discrimination can happen below the level of conscious awareness (Weiskrantz, 1990; Boyer, Harrison and Ro, 2005). The studies are typically about blindsight patients, but the crucial conclusion is that in order to categorize, and therefore sort colours and even Gabor patches, one does not need to be consciously aware of them.

Moreover, there have been a number of studies indicating that colour priming effects can take place below the threshold of awareness in neurotypical subjects (Breitmeyer, Ömen and Chen, 2004; Breitmeyer, Ro and Singhal, 2004; Breitmeyer *et al.*, 2007). Both of these considerations seem to support the thesis that abstract colour perception does not require conscious, low-level contents. The suggestion, then, is that the abstract properties of the scene, say whether there are 'many different colours', are not dependent upon the conscious representation of each of the colours that get computed over to form that abstract representation. Although the colour of each object needs to be *unconsciously* represented in order for the statistic to be computed (at least in a low-resolution format), it is much less clear the representations of each coloured object need to be *consciously* represented.[38]

A minimalist interpretation of scene statistics seems to be the most parsimonious interpretation and is roughly as follows: the information that was inputted and transformed into a summary statistic is unconscious, but the summary output, when attention is partially focused on it, is phenomenally conscious. Furthermore, the scene statistic is sufficiently low resolution, so it only consciously represents high-level properties (i.e. average colour, colour diversity, etc.) and not

[38] See Phillips (2015) for a similar minimalist interpretation of Bronfman *et al*. (2014).

information about the individual items which were used to form the statistic in the first place. There is little reason to suggest that scene statistics depend on consciously representing each item that was computed to form the statistic. It *could* be true that scene-statistical representations contain consciously represented items, but more evidence for this possibility would be needed. Merely assuming the phenomenally conscious status of the colour contents that were computed over to form the diversity judgment is problematic considering the evidence of unconscious colour processing cited above. Since it seems scene statistics are low resolution representations, Cohen, Dennett and Kanwisher (2016) appear to have the correct theory of their nature.

It should be noted that the appeal to unconscious colour processing is simply meant to block the inference from the fact that summary statistics are phenomenally conscious to the conclusion that the items that compose the statistics are likewise conscious. That is, one is not forced to believe that the individual items are conscious just because the statistic is conscious. Thus, the argument should be construed as an attempt to shift the burden of proof to those that endorse the inference. However, there is more direct empirical evidence against the inference. For instance, Ward, Bear and Scholl (2016) showed that consciously perceiving summary statistics occurs independently of consciously perceiving individual items. In those experiments, which were meant to replicate Bronfman *et al.* (2014), it was found that subjects could consciously represent summary statistics with either coarse-grained conscious access to individual items or no access at all. So, it seems that individual items are not necessarily conscious if a summary statistic is conscious, but, moreover, there is

direct evidence that individual items need not be conscious if a summary statistic is conscious.

3.2. The refrigerator-light illusion

Cohen, Dennett and Kanwisher (2016) argue that, along with the utilization of scene-statistical representations, one of the reasons people have the impression that their experience is quite rich is because eye saccades happen so quickly and effortlessly that they occur outside of one's awareness. For instance, they claim that '[t]his gives people the false impression that they perceive more than they actually do in a given instant because they are not aware of the serial manner in which they accrue information (i.e., one saccade after another)' (p. 328).

The idea is that since whatever one attends to is consciously represented with quite a bit of detail, and saccades occur rapidly and without notice, people will systematically overestimate the richness of experience at a particular moment. For instance, imagine being in a room with nine people and seeing A, B, and C (who are standing next to one another) in detail, and then D, E, and F (who are also standing next to one another) in detail, and then, finally, G, H, and I (who are standing next to one another) in detail. Eye saccades occur at a rate of about 2–3 per second, and so their prediction is that, because those shifts happen so quickly and without notice, you will take yourself to have seen them all at once. This explanation is almost certainly correct. In what follows it will be shown that the refrigerator-light illusion might actually function in several ways, however. Not only might attention shift across a scene as a result of eye saccades, but there is good reason to suppose the unconscious shifting of attention might sometimes result from an additional mechanism. First,

however, it will be argued that there are further contents that can be embedded in representations (actively held in working memory) that do not depend on the unconscious shifting of attention.

4. Additional Richness

In this section it will be argued that sparse theorists have additional tools at their disposal to explain additional richness in experience in a way fully consistent with the denial of overflow. The first option, which appeals to feature binding, straightforwardly suggests that consciousness is indeed richer than sparse theorists have assumed without assuming phenomenal consciousness overflows access. Alternatively, the second option, which appeals to latent working memory, will propose an additional way in which the *illusion* of richness can be explained.

4.1. Feature Binding

Feature binding is the process by which various forms of information about an object are integrated (Ding *et al.*, 2017). For example, the shape and colour of an object are often bound together. Feature binding allows a subject to better remember what they had seen at a previous time and seems to happen fairly automatically and from a young age (Feigenson and Halberda, 2008). This idea is reminiscent of the notion of an object-file or event-file often used in cognitive science (see Carruthers, 2015, for a discussion). The idea is that the information pertaining to some object is contained in an 'object-file' such that when the file is entered into working memory, the various conceptual and non-conceptual contents that are bound together are also

made available in working memory. For instance, the object-file for ‘tomato’ might contain the properties RED and ROUND. And when encoded in working memory one would thereby access information about, say, the colour and shape. As a result of such access, the non-conceptual content would be globally broadcast and would be conscious.

Feature binding allows information to be compressed into a manageable unit without losing much information, whereas a scene-statistical representation encodes information with a great deal of loss. There is a lot of support for this idea. For example, a 1997 experiment by Luck and Vogel showed that:

> The four-item limit was shown to apply to integrated objects, not features within objects. For example, when objects in an array of bars could differ on four dimensions (size, orientation, color, and presence or absence of a central gap), subjects could retain all four dimensions at once as easily as retaining any one. The performance function of pro- portion correct across increasing array size (i.e., increasing number of array items) was practically identical no matter how many stimulus attributes had to be attended at once. This suggested that the capacity limit should be expressed in terms of the number of integrated objects, not the number of features within objects. The objects serve as the chunks here. (Cowan, 2001)

This implies the amount of information a subject can consciously see might be larger than is often assumed.

The work of Luck and Vogel (1997), and much of the work on working memory that has since been produced, has supposed working memory is constituted by ‘slots’. That is, discrete object-files containing any number of possible features. However, this model is not without its opponents. Recently, many have pushed for an alternative way to model working memory that describes working memory as being a continuous resource (Bays and Husain, 2008; Horowitz and Cohen, 2010; Gross, 2018; Wu and Wolfe, 2018; see Ma, Husain and Bays, 2014, for a review of the

debate). The resource model suggests that working memory can be distributed flexibly across all items in a scene. According to this view, the measure of working memory is not the number of items remembered, but rather it is the precision of memory. Thus, the resource model of working memory suggests one might remember many more than 3–4 items, but they would be imprecise representations. Although the debate is still fairly lively, with proponents on either side, the slot model is utilized in this paper since it is the model that is generally invoked in the phenomenal overflow debate.[39] However, even if it turns out that the resource model is correct, that finding would be orthogonal to the arguments in this paper. That is, even if working memory is a continuous and flexible resource, it is still thought to be a *limited* resource. The claims in this paper are meant to provide a way of explaining additional riches or the intuition of richness. Thus, they do not depend on a particular view of working memory.

Before moving on, a worry will be considered.[40] There have been a number of studies that have called the original finding of Luck and Vogel (1997) into question. For instance, it was found that remembering the colour of an object does not entail that one remembers the orientation of the very same object (Fougnie and Alvarez, 2011; also see Alvarez and Cavanagh, 2004; Fougnie and Marois, 2009; Hardman and Cowan, 2015; Olson and Jiang, 2002; Schneegans and Bays, 2019, for a review). The suggestion is that there are indeed limits to the number of features one can remember. One response to this worry is that the view set out in this paper does not

[39] Although, see Gross (2018).
[40] I am grateful to two anonymous reviewers for pushing me on this point.

depend on the idea that working memory has an unlimited capacity, only that it can be increased somewhat, by feature binding. So, even if maintaining feature bindings is attentionally demanding, the argument is unaffected if more than one feature can be bound into a particular slot without affecting the content of the other slots.[41]

However, perhaps more importantly, there is indeed recent evidence for the possible increase in capacity, albeit the evidence is about familiar real-world objects rather than the somewhat simple items utilized in many experiments. Crucially, this evidence does not result from the neural signature of long-term memory, but rather from contralateral-delay activity (CDA), which is a signature of sustained working memory. Brady, Störmer and Alvarez (2016) sought to investigate the relationship between attention, working memory, and *familiarity*, using CDA as a measure. They showed two things. First, they showed that the working memory limit for real-world objects (objects with a structure) is greater than the limit for simple items like colours. Second, they showed the number of real-world objects one could hold in memory increased as a function of exposure time. Participants saw either colours or objects and saw them for either 200, 1,000, or 2,000 milliseconds each. While they performed each trial, they also performed a verbal interference task (rehearsing two digits) to ensure the use of visual memory and not verbal memory. The number of objects and objects with detail that were remembered increased with an increase of exposure time. However, the number of colours that were remembered plateaued for

[41] One plausible concern is that both sides of the overflow debate already agree on a 3–4- item limit of working memory and the position taken in this paper has done nothing to argue that more than 3–4 items can be encoded at a time. However, it should be noticed that most experiments appealed to when people critique the sparse view involve pretty simple stimuli such as numerals or individual colours. But once more complicated objects are introduced, the sparse view seems more compelling.

exposures greater than 200 ms; whereas the number of objects that were remembered and the number of detailed objects that were remembered both continued to increase as the exposure time increased (*ibid.*). This supports the idea that once you are more familiar with a particular stimulus you should be able to encode more of its features to working memory. Given the initial stimuli utilized in the Sperling and modified Sperling experiments, namely colours and numbers, it is unsurprising subjects might only see 3–4 items plus some summary statistic. However, as the evidence seems to show, in real-world scenarios when people look at a structured scene for only about two seconds, a richer and more vivid experience should be expected.

Although the authors do not think this enhanced memory capacity results from interactions between visual working memory and episodic long-term memory, they do suggest that the enhanced capacity might result from background knowledge or familiarity. Indeed, as will be illustrated in Section 5, background knowledge does seem like a plausible component of the explanation for the increased capacity to remember features of real-world objects.

So, to summarize, the relationship between feature binding and phenomenal overflow is as follows. Individual properties of an object or scene are bound together, and as a result of such binding the number of *properties* available for reasoning and reporting is increased beyond the standard 3–4-item limit. Thus, the number of phenomenally conscious properties can be greater than is assumed. So, although clearly in the purview of a sparse theory, feature binding allows the expansion of contents held in working memory. Feature binding thus seems like a plausible resource an *opponent* of overflow might utilize that would make their view more

compelling to Block or other proponents of phenomenal overflow. In other words, there is additional richness that can be explained without having to posit consciousness at the level of sensory memory.

4.2. Latent working memory

The ability to hold information in working memory, that is, the ability to encode information in an accessible state, has been assumed to require sustained activity (Cowan, 2001). Normally attention is said to explain this sustained activity. However, many have recently suggested that information can be encoded and retained in an 'activity-silent' manner. This phenomenon is known as activity-silent working memory and has been investigated by a number of psychologists (Lewis-Peacocke *et al.*, 2012; Rose *et al.*, 2016; LaRocque, Lewis-Peacock and Postle, 2014; Nikolaev and van Leeuwen, 2019; Stokes, 2015; Mongillo, Barak and Tsodyks, 2008; Wolff *et al.*, 2017; Masse *et al.*, 2019, Manohar *et al.*, 2019). Although some have called this idea into question, the dispute is merely verbal as these authors prefer to call activity-silent states a version of sensory memory rather than working memory (Trübutschek *et al.*, 2019). But it should be noted that, regardless of what one decides to call these states, the argument is unaffected. What will be clear in the following is even if activity-silent states are better thought of as sensory states (or any other form of memory that is distinct from working memory), they will not become phenomenal unless they are sustained via attention. The point will be that it is the ease with which one can attend to certain unconscious states (whether they are thought of as sensory states or unconscious working memory states) that partially explains the apparent

richness of experience. Thus, the argument does not depend on the states being characterized one way or another.

The idea of activity-silent, or latent, working memory depends on a model known as the synaptic model of working memory which suggests residual calcium deposits at synaptic terminals serve as a memory buffer with a decay time of about one second (Stokes, 2015). This information trace is consequently recoverable as a result of goal demands. Some draw analogies with long-term memory (Rose *et al.*, 2016). The upshot is the activity-silent manner in which information is held does *not* require attention and is assumed to be unconscious. However, the contents are easily recoverable by attention given task relevance and consequently can become phenomenally conscious. Although Block, and other defenders of phenomenal overflow, would of course deny that a representation merely being unattended suffices to show it is unconscious, the activity-silent states are stored without above-baseline neural activity and everyone in the debate agrees that phenomenally conscious states are associated with some neural activity. So, both sides of the debate should agree that activity-silent states are unconscious.[42]

Rose *et al.* (2016) found that activity-silent neural patterns — which correspond to representations of items not maintained by attention — could be reactivated by single-pulse transcranial magnetic stimulation (TMS) or by attention. However, they only found this facilitation effect for items that were possibly relevant for the task and they found no facilitation effect for items that were not possibly relevant. Once the task suggested that an unattended item would no longer be needed,

[42] I thank an anonymous reviewer for bringing this point to my attention.

directed TMS did not reactivate the representation. This suggests that there is a difference between activity-silent items and items that have been dropped from working memory altogether. A number of other experiments have also shown that task relevance predicts whether a representation is encoded in an activity-silent state or whether it is simply dropped from working memory altogether (Stokes, 2015). Others, however, have found that activity-silent working memory is not restricted to task relevance (Bae and Luck, 2019; Klatt *et al.*, 2020). But notice, within the context of an experiment relevance is given by the task at hand. The instructions a subject receives largely constrain what is deemed as relevant. Outside of an experimental set-up, however, when there is no task, much of the world can still be evaluated as relevant (Corbetta and Shulman, 2002).

The proposal is that subjects can have the illusion of the richness of experience in a way that does not appeal to saccades nor to scene-statistical representations. The idea that seems to emerge, given the results of these studies, is that attention can recover contents encoded in activity-silent working memory. Moreover, attention can be allocated to these contents in ways apparently unknown to subjects. This process thus seems to be a distinct sort of refrigerator-light illusion that does not *depend* on saccades. One thing to notice is there are plausible cases where saccades *might* have a role to play in this process, however. For instance, Nikolaev and van Leeuwen (2019) propose that scenes can be built up from activity-silent representation, but as the result of eye saccades. This goes a long way to explaining the refrigerator-light illusion as many, including Cohen, Dennett and Kanwisher (2016), understand it. The point of this section is not to show that saccades are

irrelevant, but rather that the refrigerator-light illusion can operate independently of eye movements.

Sergent (2018) makes a similar appeal to activity-silent representations. She emphasizes the temporal dynamics of consciousness and claims that 'although there is a limit in the amount of consciously accessed information at any given moment, information that is not initially accessed can still be at a later time' (p. 9). The idea is that unconscious representations may *become* conscious in a way that overcomes the limitations of access consciousness. The upshot is that 'despite the instantaneous limitations of conscious access, the richness of our perception might rely on flexible conscious access to a stable external world and internal buffers' (*ibid.*, p. 10). This paper is fully consistent with Sergent's proposal. The emphasis of this paper is that the flexibility of conscious access can happen in a way that eludes a subject, such that one has the illusion of additional richness. That is, accessing activity-silent representations can be swift and effortless in a way that seems best described as a variant of the refrigerator-light illusion. The upshot is, just as saccades create the illusion of richness, swift and effortless covert attention to unconscious states can also help to create the illusion of richness.

It should be noted that it is not at all obvious that activity-silent working memory representations are added to the standard 3–4 items of working memory. The claim is simply that any of the items encoded in working memory can be swiftly replaced by content encoded in an activity-silent state, such that the subject has the illusion that all of the contents were conscious at once. Just as when one unconsciously shifts one's eyes and what one then attends to seems to have been

conscious already, when one covertly shifts one's attention to activity-silent states, those states, too, seem to have been conscious already. So, activity-silent states are meant to explain the illusion of richness.

4.3. An Illustration of Additional Richness

The following is an illustration of how feature binding and activity-silent representations might operate, in addition to the usual sparse theorist's toolbox, to produce the apparent richness of experience. Consider the different visual experiences you might have while walking through a lush forest as opposed to walking down a city street or looking at a detailed and structured drawing (perhaps a blueprint). In the former case, it is plausible that you are phenomenally conscious of 3–4 items in addition to the statistical representation of 'many leaves' or 'low diversity of colour' (say, because everything is *green*-ish). Moreover, and unbeknownst to you, your eyes scan the scene and update the content that is phenomenally conscious, creating the illusion that you were conscious of everything prior to the saccade. This is the prediction of the sparse view as it is typically described. However, as has been argued, there are other tools at the sparse theorist's disposal that can explain additional apparent richness.

In the streetscape case, although scene statistics and unconscious saccades are certainly part of the explanation of the apparent richness of experience, feature binding and activity-silent representations can explain the added appearance of richness one might have. For instance, imagine walking down the city street and glancing in front of you. You are likely confronted by a plethora of structured and detailed objects. Perhaps there is a building in front of you with various details,

people walking in every direction. Moreover, there are probably storefronts, cars, and people riding bicycles. In virtue of your familiarity with many of those objects, and in virtue of your visual system's capacity to bind their features together in a way that increases representational complexity, it is plausible that you will encode many of the object's component features (colour, shape, texture, etc.). However, since there are constraints on the number of items and features you can encode at a time, many of the bound representations will fall out of conscious awareness and become activity-silent. Unbeknownst to yourself, and depending on momentary shifts in apparent relevance, you can reactivate those activity-silent representations in a way that creates the illusion that they were conscious all along. That is, you are subject to a different sort of refrigerator-light illusion. And, in virtue of the structure and familiarity of the city scene, the apparent richness of your experience can be greater than in noisy contexts. Given the structure of much of the world, the prediction should be that people can have a richer experience than is typically assumed.

Of course, one might fail to bind many of the scene's features together because one's attentional resources have been exhausted, or because attention was not directed at the various stimuli for long enough. In those cases, the explanation for an increase in the apparent richness of one's experience might simply boil down to the constant shifting of covert attention to latent representations. So, the processes that are part of the sparse theorist's toolbox might disassociate; however, the processes might work in tandem as well.

5. Conclusion

The phenomenal overflow debate is concerned largely with the subjective impression of consciously perceiving a rich and detailed scene. The most promising explanation of this to date is dependent upon statistical representations and the unawareness of saccades. Although these processes are indeed effective explanatory tools, there seem to be other resources available to sparse theorists that might explain a higher degree of richness when there is a sufficiently salient structure in the environment; for instance, a city streetscape or a detailed drawing (perhaps a blueprint or some other visual representation of structure). As mentioned above, both are organized wholes with many different component parts, all working together. Rather than having to keep in mind the various individual features one by one, one might instead organize them into groups. In fact, some of the work is already done since both are typically arranged in a way that makes a certain structure apparent. A plausible reason why city streetscapes or blueprints don't appear as a noisy mess is because people (especially those familiar with the components of the stimuli) have abilities to parse and structure the information into usable units. In forest scenes, on the other hand, there is, in fact, not enough structure for observers to exploit. Importantly, assuming a global broadcast view of consciousness, if the features that are integrated are encoded in working memory, then they are conscious. Moreover, if they are held in so-called unconscious working memory, then they are available for attention in a quick and effortless way. This is not to say scene statistics are never utilized — in fact they are most certainly always employed by the visual system. The upshot is that, in addition to scene statistics and unconscious saccades, there are other

tools at the sparse theorist's disposal that plausibly explain the intuition that one's experiences outstrip that which one can remember.

Chapter 4:

The Bounds of Experience: Representationalism and the Laws of Appearance

1. Introduction

It is generally accepted that the scope of what we can represent outstrips the scope of what exists. (See Brentano, 1874/1973.) We can represent ghosts, ghouls, and goblins, even though none exist. We can also perceptually represent such things in imagination and in dreams. Additionally, we can also represent apparently incompatible states of affairs. However, we cannot, it seems, perceptually represent such states. For instance, one cannot perceptually represent pure-blue and pure-green at the same location and time even though we can represent such an occurrence in thought.[43] Moreover, one cannot perceptually represent that an object has a color without also perceptually representing it as taking up space, even though, in principle, one could represent that content as well.[44] The reason is that there are constraints on how things could seem; these are the laws of appearance (Pautz, 2017, 2020; Tye, 2020; Speaks, 2017; Lycan, 2019, Sainsbury, 2022).

These constraints present a problem for strong-representationalism ("representationalism" from now on) which maintains that having an experience is

[43] Now, there are likely going to be constraints on *belief.* So, if one is an inferentialist about belief, for example, then the possible contents that one can believe or endorse will partly be a function of the entailment relations between the components of the representation. Even if one could represent a contradiction, one might not be able to *believe* it. This paper is thus contrasting perception with *representation,* not with belief.

[44] There are other examples as well (many of them listed at the outset of Section 2), but we shall restrict our discussion to a select few since they are representative of a general problem.

nothing over and above representing the world to be a certain way. Since there are no laws of representation in general, we shouldn't expect there to be any laws of experience, but it seems that there are.

The goal of this paper is twofold. First, I will present a series of considerations that put pressure on Pautz' particular argument for the metaphysical necessity of the laws of appearance. I argue that his particular methodology is weak and doesn't justify *belief* in the metaphysical necessity of the laws of appearance. Second, I argue that the only plausible account of the metaphysical necessity of the laws does not threaten representationalism about consciousness either. The upshot is that the truth of representationalism about consciousness is not threatened.

This paper proceeds as follows. In section 2, I present the laws and discuss their supposed metaphysical status. I then make explicit their apparent incompatibility with strong representationalism. In section 3, I present a series of considerations that raise doubt for the claim that the laws of appearance are metaphysically necessary. Specifically, I argue that the laws only seem necessary because of facts about imagination. Importantly, this is consistent with physicalist versions of representationalism. In section 4, I then present an argument taken from Sainsbury (2022) showing that there is a plausible, but narrower, version of the laws that might be metaphysically necessary. I conclude that either way, the laws of appearance do not undermine physicalist versions of representationalism. In section 5, I sum up.

In what follows, I assume, rather than argue, that representationalism should be thought of as a reductive theory of consciousness. Much of what I say only gains traction if we understand representational states to be, at the very least, token

identical to or grounded in some underlying physical state. The conclusions I draw might not be compatible with non-reductive versions of representationalism about consciousness, such as Chalmers (2004), but so be it.

2. The Laws of Appearance

Taken from Pautz (2020), the laws are as follows[45]:

1. *Exclusion law:* You cannot experientially represent the same surface as pure-blue and pure-green at the same time.
2. *Berkeley's law:* You cannot experientially represent that something has a color quality without also experientially representing that it takes up space.
3. *No Logical Structure:* An experience cannot have as its *only* representational content: there *is either red square in front of me or green sphere on the right.*
4. *Perspectival law:* An individual cannot experientially represent merely *that there is a cube somewhere in reality*, without any "perspectival information" about its location and apparent shape from "here".
5. *No High-Level law:* There cannot be an experience whose *only* content is a free-floating "high-level" content, like *there's a tomato in front of me*, without having *any* contents attributing lower-level features like shape or color.

For brevity only the first two alleged laws will be discussed. My discussion is decidedly restricted since arguing against all of the laws is beyond the scope of this paper. But the methodology I employ to undermine the belief in the metaphysically necessity of the first two laws should provide us some insight into how to undermine

[45] The following might not exhaust all the laws.

our commitment to the remaining ones. Indeed, the diagnosis of the metaphysical status of the laws offered in section 3 makes this point clear since it suggests that the best explanation of the laws derives from facts about the visual system, and not the supposed, yet nebulous, capacity of rational insight into the metaphysical constraints on experience. So, very generally, I suggest that one should look at the science, and not our intuitions, in order to argue against the metaphysical necessity of the laws.

Importantly, this paper specifically engages with Pautz's (2020) defense of the laws; that they depend on their intuitive plausibility. Pautz's tactic is to show that that intuition can justify claims about the metaphysical nature of experience. Consequently, this argument will be the target of the present piece. There may be other ways to defend the metaphysical necessity of the laws but I'm unsure how successful they could be. As we will see, however, Sainsbury (2022) defends the necessity of the laws by appealing to human vision, but the scope of his defense is decidedly narrower than what Pautz has in mind. In fact, all Sainsbury defends is that if one is a normal human, then the laws hold with metaphysical necessity. Pautz's claim, in contrast, is that the laws are necessary in the sense that they constrain all *possible* creatures (including, of course, *abnormal* humans). We will come back to this issue in section 3.2.

According to Pautz, it is *intuitive* that each law holds for every conscious experience and for every conscious experiencer. That is, he thinks that it is intrinsically plausible that there is no conscious experience or experiencer that enjoys the freedom that might be expected if one thought that the content of experience really does depend solely on a representation's content.

Supposing that the alleged laws of appearance are genuine laws, there are a couple of things one could think about them. On the one hand, one might think at least some of the laws are metaphysically necessary (perhaps even all are). That is, some of the alleged laws hold for all possible experience and *experiencers*, and in all possible worlds. On the other hand, given certain assumptions about representationalism (as opposed to weak intentionalism[46]) about conscious experience, one might think all the laws are merely contingent. That is, all of the laws hold in most cases, but not all. Perhaps there is a class of aliens or even humans in unusual circumstances, that violate the laws. These two interpretations are mutually exclusive and ultimately lead to a puzzle that Pautz refers to as the puzzle of the Laws of Appearance.[47]

The first thesis is Necessitism: *some* of the laws are metaphysically necessary. The second thesis is Contingentism: *all* of the laws are contingent special science laws that can be violated by alien minds or humans in unusual circumstances. (Pautz, 2020, p.258) In order to argue for Necessitism, what one needs to show is that some of the laws hold in all worlds and for all possible experiencers. However, if true, this thesis is puzzling since if some of the laws are necessary in this sense, there ought to be a general explanation or else they are all brute facts (Pautz, 2020, p.259)[48]. Moreover, it is difficult to see why there should be such strict laws given the truth of

[46] "Weak intentionalism" is simply the view that all experience has a representational component. It is therefore consistent with the addition of qualia or *mental paint*. In fact, a variant of sense-data theory (sensory core theory) claims that all experiences have a representational component *as well* a sense datum component. Strong representationalism, in contrast, is the view that an experience is fully exhausted by its representational component.

[47] This puzzle is to be distinguished from the puzzle that motivates this paper, namely that representationalism is not consistent with the laws of appearance, yet the laws are intuitively plausible.

[48] Again, if they are necessary in the relevant sense, the account I provide in section 3 plausibly establishes the unity of the laws.

representationalism. In contrast, to argue for Contingentism, one would need to show that *all* the laws *could* be violated by some *possible* creature. However, this thesis is also problematic since there do not appear to be counterexamples to the laws, and moreover, it seems intuitive that there could not be any such counterexamples.[49] So, both options are problematic. This is the puzzle that motivates Pautz. I won't be engaging specifically with this puzzle, rather, I will be undermining an important reason for taking Necessitism seriously, namely that the laws of appearance are intuitively metaphysically necessary. That is, I will argue against the idea that we have rational insight into the real nature of experience.

So, one reason to favor Necessitism (though perhaps not the only reason, as we will see), is that the laws are intuitively plausible when taken to be necessary. (Pautz, 2020, p. 259) One way to motivate taking this intuition seriously is the following argument: we agree that it is metaphysically necessary that a blue experience resembles[50] a purple experience more than an orange experience, and we agree because it is intuitively plausible. Likewise, the Exclusion law is intuitively plausible. And, as Pautz claims, by parity, if we accept claims about similarity of experience, we ought to accept claims about necessity of experience.

Now, one might question this argument from the get-go, however. The reason is that facts about phenomenal similarity are quite different from facts about metaphysical necessity. It is plausible that similarity judgments might be justified

[49] On first glance it might seem strange to take seriously our intuitions regarding merely possible creatures or merely possible states of affairs, but there are examples of this sort of methodology in the literature, (see BonJour, 1998).

[50] A note: resemblance relations amongst colors are no doubt built into the human visual system. Whether or not these relations are metaphysically necessary is a further step that is not argued for by Pautz.

first-personally, since the properties that give rise to judgements about phenomenal similarity are themselves first-personally accessible. Moreover, the similarity space in which the properties are located is at least plausibly first-personally accessible too since one can *just see* that some phenomenal properties cluster in a way that is related to their similarity. That is, one can just see that phenomenal red is more like phenomenal orange than phenomenal blue; we don't have to make inferences about their similarity. On the other hand, facts about metaphysical possibility do not seem to be phenomenally accessible since under a reasonable understanding of phenomenal properties—that they are nonconceptual in nature—the property of *being possible* is not itself phenomenal.[51] Although I take this to be a powerful objection, in what follows I aim to show that even if we allow that the laws strike us as necessary, representationalism can still be defended. So, for the sake of argument, I will put these considerations to the side and grant Pautz's claim that the laws do *seem* metaphysically necessary, whatever that is supposed to mean.

As I said at the outset, the truth of the laws poses a problem for representationalism because there are no such laws for ordinary representational states. However, the truth of the Necessitist thesis poses an even more striking problem for representationalism than the weaker contingentist thesis. The problem for representationalism raised by the Necessitist interpretation is that unless, one thinks that one can only represent that which exists, the possible content expressed by

[51] Now, one might object that *possibility* might be constitutive of one's experience. For example, perceptual experience might represent the various ways one might interact with their environment, or it might represent possible consequences of interacting with one's environment (e.g., falling off of the side of a cliff). However, even if one thought that, it is a much stronger claim to suggest that *metaphysical* possibilities are phenomenally conscious. This is the notion of possibility that is not obviously constitutive of one's experience.

representational states is not obviously bounded. The reason is that the standard account of representationalism about experience claims that existence does not determine content. As we saw above, one can represent ghosts, ghouls, and goblins even though none exist. As Pautz puts it, most representationalists are committed to this sort of "existence neutrality". This suggests that if a representationalist is going to maintain that experiential content supervenes on representational content, then, at most, they should be committed to Contingentism about the laws.[52]

3. In Defense of Representationalism

In this section I defend the representationalist thesis in two ways. First, I present a debunking argument against the metaphysical necessity of the laws. I argue that the laws only *seem* necessary because of contingent facts about human vision, thus undermining the idea that we have rational insight into the real nature of experience. Specifically, I suggest that the best explanation of the intuition of metaphysical necessity is that our mind/brains are structured in such a way that excludes certain sorts of experiences, and not that we have the capacity to know the metaphysically necessary truths. Importantly, however, since most

[52] Now, Contingentism might also spell trouble for the representationalist, but the problem is decidedly less striking. The contingentist thesis puts pressure on representationalism since there are no obvious constraints on representation that are even nomologically necessary, let alone metaphysically necessary. So even if the weaker contingentist thesis were true, the representationalist still owes us an account of the distinction between representation and phenomenal consciousness. However, since there are physical constraints on the way phenomenal consciousness is realized in the brain, we could explain the nomological necessity of laws in virtue of those structures or properties. Importantly, this sort of explanation is fully within the purview of a standard physicalist version of representationalism even if it, strictly speaking, adds more to the thesis than as I have defined it above.

representationalists are also physicalists, one can simply claim that the reason experience is constrained differently than thought is that there is a distinct and contingent physical realizer for experience than for thought. Although this is strictly speaking more than the traditional representationalist thesis claims, it does not require a shift to an alternative theory to accommodate it. One can maintain one's representationalist credentials even if one claims that experience is distinct from thought in that it depends on distinct neurocomputational processes, structures, or properties. The argument can be understood thus:

1. There are properties of perceptual representations that explain why we cannot perceptually represent violations of the laws of appearance.

2. Since there is significant overlap between properties that explain perception and those that explain imagination (e.g., imagination utilizes the resources of perceptual systems), imagination is subject to the very same constraints.

3. This in turn explains the intuition that laws of appearance are metaphysically necessary.

4. Therefore, there is no other reason to think the laws are, in fact, metaphysically necessary.

I defend each premise in what follows.

I then suggest, in accordance with Sainsbury (2022), that there is a plausible, albeit, uninteresting, version of the laws, that we might have good reason to regard as metaphysically necessary. For instance, one suggestion, considering the argument that intuition is not a good guide to metaphysical truth about possible forms of experience, would be to restrict the predicate "conscious visual experience" to normal human

vision. This is the tactic utilized by Sainsbury (2022). The idea is that if "conscious visual experience" picks out normal human vision, then Necessitism is true. After all, one would have to be quite abnormal to violate the laws. However, I argue that if one understands the predicate "conscious visual experience" in this extremely gerrymandered way, representationalism is not threatened either. I discuss each option in turn.

To begin our discussion, I want to draw attention to the fact that defending the laws by appealing to the first-person perspective seems to beg the question against the reductive representationalist. After all, any plausible reductive representational thesis of consciousness is committed to the idea that we can fully explain phenomenal consciousness using the terms and predicates of some naturalistic language (the language of cognitive psychology or whatever). Importantly, the terms and predicates of all naturalistic idioms are typically thought to be fundamentally third-personal.[53] As a result, representationalism, understood as a reductive theory of experience, makes certain commitments to the nature of phenomenal experience.

Consider global workspace theory, for example, though I am not committed to this view in particular.[54] It claims that phenomenal consciousness is constituted by nonconceptual content that is made available to several cognitive subsystems, namely, the systems responsible for decision making, reflection, reasoning,

[53] One might insist that science does depend on *some* indexical content. For instance, when one refers to the experiment that they did *last week*, or to the one that they are *presently* engaged in, or to the information presented on *that* scanner, one is utilizing first-personal content. Although there is a huge literature on indexical content that I could not possibly do justice to at present, none of the previous indexical seem to be *essentially* first personal. For example, demonstratives can plausibly be given an analysis in terms of space/time coordinates that do not make irreducible reference to one's perspective.
[54] Indeed, one might instead endorse some neighboring theories without the neurological commitments of the global workspace theory. For instance, one might instead opt for Prinz's (2012) AIR theory, or Tye's (1995) PANIC theory. What I say is neutral between these theses.

remembering, and reporting. (Block, 2008; Baars, 1998; Dehaene, 2014; Carruthers, 2019.) Moreover, those states are composed of representations with nonconceptual content. (See Beck, 2019 for a plausible account of nonconceptual content based in Weber's law.) So, the implication is that phenomenally conscious states are nothing over-and-above access-conscious states with nonconceptual content. The upshot is that token phenomenal properties are simply *identical* to token globally broadcast nonconceptual contents. A token phenomenal property, then, is identical to a token representational-functional property, and ultimately a physical one. If one successfully provides such an account, then it seems likely that one would only avoid begging the question against reductive representationalism if one expressed the laws in terms that are consonant with that account.

Take the exclusion law for example.

Recall:

Exclusion law: You cannot experientially represent the same surface as pure-blue and pure-green at the same time.[55]

Exclusion law*: Pure-blue nonconceptual content that is ascribed to a certain surface S cannot be globally broadcast at the same time as pure-green nonconceptual content ascribed to that same S.

The recast law claims that pure-blue nonconceptual content and pure-green nonconceptual content cannot, given the way our visual system works, be ascribed to the same surface at the same time. The resultant question is about modal status of

[55] One might insist that Hardin (1998) casts doubt on this law. However, it should be noted that Hardin simply showed that people can sometimes be made to see reddish/green. The exclusion law has to do with the experience of pure colors, not blended ones.

such a law. If it is a special science law, it is at least metaphysically contingent (perhaps nomologically necessary).[56] That much seems uncontroversial. The difficult question is whether it is metaphysically necessary. Remember, the main reason for thinking the exclusion law is necessary is said to be that it is intuitively plausible. Once the laws are framed in terms acceptable to an actual representationalist about experience—by invoking third-personal terms, rather than first-personal terms—the intuition of metaphysical necessity is not plausible.[57] This is uncontentious since, even according to Pautz, there do not seem to be constraints on representations, as such. Recasting the laws in purely representational terms, then, should not lead to the intuition that the recast law is metaphysically necessary, if one doesn't also think that there are constraints on representation proper, which Pautz doesn't.

On the other hand, Pautz might simply reply that *I* am begging the question in favor of representationalism; the thesis which is in dispute. After all, part of the reason I would be justified couching the laws in terms of reductive representationalism, is if there were some antecedent reasons to believe it. But again, that is what is being disputed. So, there seems to be a standoff. My claim is that from the third-person perspective, the laws do not appear to hold with metaphysical

[56] I'm understanding metaphysical possibility to outstrip physical or nomological possibility. The idea is that metaphysically possible states of affairs supervene on metaphysically possible worlds, whereas nomologically possible states supervene on nomologically possible worlds, and, importantly, the set of metaphysically possible worlds is larger than the set of physically possible worlds. This does not strike me as a particularly contentious assumption since it is standard to assume that the laws of physics could've been different.

[57] One might insist that introspection, at least in humans, represents experience *as experience*, thus not all representations can be given in purely third-personal terms. However, when one introspects on one's experience as an experience, they presumably need to bring to bear the concept EXPERIENCE. Unless one thinks that that concept is essentially an indexical concept (which would be weird since we can apply the concept to others), then it's unclear that we cannot give a fully third-personal account of it.

necessity, whereas Pautz claims that given certain facts about the first-person perspective, reductive representationalism is not plausible.

Okay, then how shall we proceed? I think the appropriate way forward is to present a challenge to Pautz analogous to the challenge presented to moral realists. Namely, we should ask what could justify one to believe that one has access to deep facts about the nature of subjective experience given that our capacities to form judgments are "saturated with evolutionary influence" as it were. (See Street, 2006, p. 114) And, perhaps more to the point, why think we should have access to deep facts about the nature of *possible* experience.[58] Evolution shaped our dispositions to form judgments generally, and judgments about experience are just one example. The problem for moral realists, and I think, for Pautz also, is that there is no reason to believe that the judgments we form about such matters are truth-tracking rather than simply conducive to survival. Indeed, even if they happen to be true, it could just be a coincidence. If it is a mere coincidence, then it requires further argument to show that facts about metaphysically necessary laws *explain* our intuitions about them. So, Pautz owes us an explanation of why, given the fact that our judgments needn't be truth tracking, our intuitions about the real metaphysical nature of subjective experience are reliable. Of course, his claim seems to be that we have a capacity for metaphysical insight, but again, why should we think that our judgments are justified given the fact that evolution presumably couldn't care less about judgments of that

[58] It is true that one might have access to deep facts about math, for example, but access to these facts is presumably explained by more primitive capacities. Gaining access to deep facts about *possible* experience seem much less plausible. Now, it is logically possible that we have an accidental capacity to access possible experience, but it seems that the burden is on the proponent of such a view to make the case. Simply putting forward intuitions is not obviously the correct strategy.

kind? My contention is that we have *no* such capacity, and that our intuitions about the scope of subjective experience is explained by facts about perception and imagination. Taken together, this provides a powerful reason to be skeptical of the laws understood as metaphysically necessary. In the next section I make this idea explicit and defend it.

3.1. The Constraints on Perceptual Representation

In this section I suggest that there are plausible constraints on perception which, as I argue in the subsection that immediately follows, apply to states of imagination as well. I present two very different accounts, but they both show roughly the same thing. One account of visual representation comes from Tye (2020) which suggests that visual representations have the structure of an array, as it is understood in computer science. According to this view, mental images (whether they are perceptions or imaginings) have parts that stand for parts of the visual scene, much like an actual image (*the part principle*). Importantly for Tye, the cells that compose the array each contain a symbol that stands for an environmental simple (or atom). So, the reason the Exclusion law holds is that each cell can contain only a single-color symbol (say, blue) and not more. Importantly, if perceptual representation has this sort of format, and if imagination shares resources with perception (which I will show in the subsection that follows), then whatever one says about the constraints on perception, one should also say about the constraints on imagination.

An alternative format-based account comes from Beck (2018, 2019). Beck argues that the mind traffics in analog representations; as opposed to the digital,

"array format". There are many views about the distinction between analog and digital representations, but Beck's view accords with the *mirroring conception*. The idea is that a representation is analog in virtue of being structurally isomorphic to, or otherwise mirroring, what it represents (Beck, 2018, p.2.) For example, a mercury thermometer is an analog representation of temperature because the height of the mercury mirrors the temperature. Although there has been much debate about the appropriate definition of *analog*, this debate does not concern me here. The point worth emphasizing is that it is plausible that constraints on perceptual representation fall out of the geometric structure of analog representations, understood in the mirroring sense. For instance, in virtue of the fact that the mercury cannot represent two distinct temperatures, since the mercury cannot have two incompatible heights at once (a physical fact that I understand to be uncontroversial), the exclusion law about the representation of temperature seems to follow naturally. If it is the case that color, for instance, is likewise represented in such a fashion, then it is natural to conclude that the exclusion law follows from the analog format of color representation. There are many other theories that likewise claim that color contents fall out of the geometric structure of mental representations or the spaces that they represent (c.f. Lee 2021; Lande 2018; and Lande 2019, for some examples.)

It should be noted that some of those discussions do make certain modal claims about representation, but they are decidedly weaker than what Pautz claims. For instance, Lande (2021) claims that, "Structural claims purport to explain modal nomological facts about the *psychological possibility* (or impossibility) of various co-occurrences of mental states by reference to a narrower notion of *structural*

possibility. For example, suppose that it is psychologically impossible to be in a type of visual perceptual state that represents the shape of a mountain ridge without also being in distinct types of visual states that represent the shapes of each peak on the ridge. It is outside the scope of this paper to say how observational data can support these kinds of nomological claims in psychology." (p.651) What is particularly striking about this passage is his skepticism of making sweeping claims about nomological possibility, let alone, *metaphysical* possibility. If facts about visual format cannot straightforwardly explain nomological possibility, then surely, they cannot ground metaphysical possibility either.

Again, if perception has this sort of format, and if imagination is a form of top-down perception (which I show in the next subsection), then the same constraints apply to both perception and imagination. So, there are facts about human visual experience that explain why one does not have perceptions which violate the exclusion law. Perhaps these are facts about digital or analog format of perception. It will be my contention that these facts then partly explain the intuition of the metaphysical necessity of the laws. What I still need to show is that imagination uses many of the same resources as perception. If successful, then taken together this provides one with reason to be skeptical regarding the idea that intuition gives us a grasp of the metaphysical necessity of the laws.

Let me briefly discuss Berkeley's law as well. Recall that Berkeley's law states that you cannot experientially represent that *something* has a color quality without also experientially representing that it takes up space. This law is deeply intuitive insofar as it seems impossible to imagine a violation of it. However, again,

some remarks about the human visual system will provide some reason to be skeptical that our intuitions support the metaphysical necessity of the law.

Feature-binding is a perceptual phenomenon whereby information about an object's form, color, texture, and location in the visual field are bound together. (Treisman, 1998.) Importantly for our discussion, feature-binding is automatic, albeit attention dependent, and occurs prior to conscious awareness of one's surrounds. For instance, it is not the case that one has a conscious experience of color separately from a conscious experience of form and then consequently combines them into a single package (what is known as an object-file in the literature). This sort of combinatorial process is what happens, however, albeit, below the threshold of conscious awareness. The reason that feature-binding provides an explanation of Berkeley's law is that, at least according to Treisman's Feature Integration theory (Treisman and Gelade, 1980; Treisman and Sato, 1990; Treisman and Zhang, 2006), our brains represent a "master map" of locations and then "[a]ttention selects all the features associated with a particular location, and works as glue to bind those features." (Bharti, et al., 2020) The idea is that spatial information is fundamental, attentionally speaking, and so when the visual system attributes certain visual surface features, such as color, it simultaneously attributes the property to a spatially located object or region.[59] And for some object to be spatially located, it must have an extension since locating an object requires placing it some sort of map (whether

[59] Now, there may be many psychologically possible spaces, some of which might not be Euclidean. This is orthogonal to the present point, however, since even if an ordinary human were to represent hyperbolic space in perception, they would presumably still, other things being equal, attribute properties to points in that space,

allocentric or egocentric). Now, although it must have some shape or other, the shape might be ambiguous.

Again, if perception is partly explained by feature-binding, then our perceptual states won't violate Berkeley's law. One question that looms in the background though is why perception *must* function like this? Isn't it at least conceivable that color properties could be attributed to a single point on a graph, thus undermining the idea that color must have form? I think it is. Perhaps in us, and in those creatures for whom feature binding functions in *this* way, it is *nomologically* impossible to represent color without representing extension. That is, perhaps given the precise structure of our visual representations, a node represents a region of space, and not a single point. But that is a far cry from the thesis that it is metaphysically impossible. Surely, we could imagine a process where color is attributed to a single point on a graph, though we'd probably *not* be imagining this from our first-person perspective.

3.2. Imagination as Top-Down Perception

So far, I have assumed that there is some naturalistic theory of conscious experience and the representations that explain it and that it can at least explain the nomological necessity of the laws (again, everyone in the debate can accept this). My preferred account is the Global Neuronal Workspace theory that takes phenomenal content to be analog in nature, but nothing hinges on that. I will also take for granted that imagination and perception share resources. I will provide some examples, but I won't argue for the truth of the examples since it is common in the literature to take them as establishing the integral relation between perception and imagination. But the

upshot is that in virtue of this sharing of resources, imagination will have some of same constraints as perception.

There is extensive evidence supporting the hypothesis that imagination is a form of perception that occurs in the absence of perceptual stimuli.[60] Not only is it functionally analogous in the sense that when one imagines some percept or perceives it, one can form judgments about it and manipulate it in one's "mind's eye", imagination also shares many of the same underlying mechanisms as perception. (Kosslyn, 1994; Kosslyn et al., 1995; Kosslyn et al., 1997; O'Craven & Kanwisher, 2000; Reddy et al., 2010; Stokes et al., 2009; Tootell, 1998; MacEvoy & Epstein, 2007, Cattaneo et al., 2012. See Cichy, 2012 and Brogaard and Gatzia, 2012 for overviews). Early studies of visual imagery showed that transcranial magnetic stimulation (TMS) to area A17 disrupted both perception and imagery of stripes (Kosslyn, et al., 1999) suggesting that the representations utilized in imagination and perception are, at the very least, coextensive.[61] More recent evidence also shows similar results. For instance, Cattaneo et al., (2012) showed that TMS directed to V1/V2 impaired the ability to encode perceptual information into working memory and to imagine the same content. Importantly, for the so-called Exclusion Law, Cichy et al., (2012) showed that imagery and perception encode not only the category of the

[60] Like many of the concepts found in the psychological literature, "perception" is not supposed to be factive. That is, like having a "memory", which does not depend on the truth of its content, one might have a "perception", independent of the correctness of its content. It is an unfortunate terminological confusion. When I, and others, use the term "top-down perception", all that is meant is that the processes responsible for perception (in the successful case) are also utilized.

[61] Transcranial magnetic stimulation (TMS) is a form of brain stimulation that uses a magnetic to produce electrical stimulation of the brain. TMS introduces noise in neural processing by inducing neuronal activity that is unrelated to the task-related activity of those neurons (Harris et al., 2008). This is important for the above studies because TMS is used to prevent the typical functioning of the brain areas directly impacted by TMS.

object perceived, but also its location. They found that imagery and perception encode location information in lateral regions of high-level visual cortex. The latter bit of evidence is telling because goes some way towards explaining why we cannot imagine overlapping objects (or properties) since imagined location information is encoded in the same way as perceptual information, which does not violate the exclusion law, at least contingently.

Crucially, since perception and imagination share resources in at least the ways just mentioned, what we can imagine is going to be limited in the same way that perception is limited. And since imagination is restricted in this sense, it can plausibly explain why we intuit the laws to be metaphysically necessary (since it explains why we cannot imagine violations of the laws). But because the systematic failure to imagine counterexamples to the laws plausibly explains why we *believe* that the laws of appearance are metaphysically necessary, we do not need to claim that the laws really *are* necessary. Finally, and more to the point, if the laws only *seem* necessary, then we have no reason to think that we have rational insight into the structure of experience. The consequence of this line of reasoning then, it that representationalism is not under threat.

If the aforementioned accounts of the format of perceptual representation or the processes and mechanisms that explain perception and imagination are correct, it seems most natural to conceive of these as metaphysically contingent. The reason is that nothing about the nature of perceptual representation or feature-binding suggests that it should be metaphysically *impossible* for there to be other creatures that are not

so constrained. Perhaps there is another world, or perhaps a creature or human in this world, that is not constrained in such a way.

Yet, the Necessitist insists on the claim that it is metaphysically impossible to violate the laws. It is my contention that the Necessitist intuition is partly driven by the very same mechanisms and properties that explain human vision and partly driven by a failure to override these failures of imagination in other ways. The intuition that it is impossible for the laws to be violated is partly explained by our failure to imagine counterexamples. Importantly, it is a problem with imagin*ability* in addition to our failure to come up with alternative reasons to undermine our intuition. So, unimaginability, along with a failure to come up with an alternative route to undermining that belief, provides the basis for the default belief that the unimagined thing or state is impossible. That is, *prima facie*, if some state is unimaginable, we won't believe it is possible. This point will be clarified below.

3.3. On the Relation between Imagination, Intuition, and Belief

So far, I have defended two premises of my overall argument. I showed that there are

constraints on perception that accord with the laws (albeit in a way that does *not* suggest the restrictions are metaphysically necessary rather than merely nomological necessary). I also argued that imagination is a top-down form of perception, thus showing that imagination is constrained in at least the same ways. I have not yet argued for the claim that the constraints on our imaginative capacities *explain* the intuition that the laws of appearance are metaphysically necessary. The

purpose of this section, then, it to defend that claim. I defend it by engaging with an objection that suggests that there are lots of unimaginable experiences that we do *not* think are impossible.

Indeed, Pautz presents this objection in his 2020 piece. He points out that there are lots of experiences that we cannot imagine, but that we do not *believe* are impossible. If there are such examples of unimaginable but conceivable experiences, then unimaginability is insufficient to explain the intuition that such experiences are indeed impossible. As Pautz put it,

> "[t]o clarify, the reason for thinking that [Exclusion] is metaphysically necessary is not the following argument: we cannot imagine counterexamples to [Exclusion] from the first-person, therefore counterexamples are metaphysically impossible. That is a bad argument. We also cannot imagine the alien color experiences of pigeons or the alien echolocation experiences of bats, but this doesn't mean that they are impossible. Rather, the reason for thinking that [Exclusion] is metaphysically necessary is just that this is intrinsically plausible." (p.259)

The same reasoning applies to the case of 1000-sided figures (Pautz, in correspondence). We cannot imagine an experience of a 1000-sided figure, but we do not thereby consider such an experience impossible. So, the idea is that there are plenty experiences that we cannot imagine, but that we think are conceivable.

Consider our commitments to actual *alien* experiences first. For instance, we all agree that bats experience the world partly through echolocation. Moreover, there are many creatures that experience colors that we do not experience, such as some birds and mantis shrimp. So again, there are a set of experiences that we do not believe are impossible, but that are not imaginable.

One way to respond to this argument is to point out that the commonsense explanation of why we can't imagine an experience of colors outside of the visual

spectrum or imagine echolocatory experiences is that, we have not had such perceptions in the first place. Obviously, we cannot imagine an experience we cannot have, given the structure of our perceptual system, so we have an explanation of our failure to imagine echolocatory experiences. Most people, I think, would claim that were we to have the appropriate physical apparatuses, we would indeed have such experiences, and thus be able to imagine them. Importantly, we can move way from common sense and can partly explain this difference in physiology in terms of selection pressure, along, of course, with other contingencies of the bat's phylogeny. Bats have distinct neural structures which explain their capacity to echolocate, which we don't have, and mantis shrimp have distinct neural structures which explain their capacity to see "invisible" colors. Surely, if we had such a structure, or genetic history, then we'd be able to imagine echolocatory experiences. So again, we have an additional reason for believing in the possibility of experiences that we cannot imagine. This is consistent with the default assumption that unimaginability entails belief in metaphysical impossibility. We just override it when we reflect on the conditions required for imagination, namely, having the perception in the first place or the right physical makeup.

Now, let's discuss another set of possible counterexamples. One might insist that there are many possible human-*like* experiences that one cannot imagine: there are 1000-sided figure experiences, 1001-sided figure experiences, 1002-sided figure experience, and so on. *Prima facie* this also puts pressure on my preferred explanation of the Necessitist intuition since, again, I argued that a failure to imagine some experience entails the belief that that experience is impossible. Since we

certainly cannot imagine 1000-sided experiences, we should also believe that they are impossible. But we don't.

Now, I do not think that there are *ordinary* human experiences with the above content (this sort of experience is then an intermediary between human and alien experience). This is made plausible given our understanding of how we represent numerosity, as an analog magnitude and not as a discrete point on a number line. Obviously, one could tell that a 1000-sided figure has more sides than a 20-sided figure, but that does not entail that one is representing it *as a 1000-sided figure*. The property of *having 1000 sides,* is not itself a phenomenal property that normal humans are acquainted with. And thus, it is not part of one's experience even if one's experience refers to an object that has 1000 sides. So, even if one might conceptually represent that a particular object has 1000 sides (because one counted or knew in some other way), it does not follow that the property of *having 1000* sides is phenomenally conscious. One might instead simply represent the 1000-sided object as "having many sides". And surely one can imagine *that* sort of experience. So, even if there are experiences of 1000 sided figure in a purely referential sense, human's do not have phenomenally conscious experiences with the content *1000 sides.*

One should grant however that it is in fact *metaphysically* possible to have an experience that represents an object *as having 1000 sides*. Of course, introspectively, I cannot have, or indeed, cannot imagine, such an experience. But one might reasonably insist that there's no reason in principle that some creature or other couldn't have a mind capable of experiencing objects *as having 1000 sides*. Of course, normal humans can't, but that's best explained by the limitations of our

minds. Now, even if it is impossible for *us* to imagine such an experience, we have some idea of what background conditions should hold for there to be a creature with such an experience. Surely such an experience would just be a more elaborate version of what we have. For instance, when presented with a couple of items, we can just see that there are 2, for instance; we don't need to count them. This is a capacity known as "subitizing" (Kaufman, et al.,1949). It makes sense that our capacity to immediately see number is limited since it would have been implausible that seeing more than a few items in a single glance would matter from an evolutionary perspective. A natural thought, however, is that there could have been other creatures in evolutionary history for whom it would matter that they represent an object as having 1000 sides and not merely 999 sides. That would surely be an unusual circumstance, but there's no reason to think it's impossible. One implication of this is that my failure to imagine such an experience doesn't cause me to believe that such an experience is metaphysically impossible since I can understand what it would take for such an experience to obtain. Without such an understanding, my own view is that we would believe it to be impossible. My claim is *not* that one needs to know about cognitive science. Rather, a moment's reflection on our own capacities should make this possibility salient: we start with introspection and notice that we can have experiences as of a 3-sided object, for that is within the normal range of our abilities, then we notice that 4-sided objects are also plausibly immediately noticed *as 4-sided objects*. Now, again, I cannot imagine a 999-sided object, but it is reasonable to assume that imagining that sort of experience isn't a wholly distinct capacity, but rather a more complicated version of what I have.

In each of the cases above, the upshot is that we have a default assumption that the failure to imagine some state entails having the belief that it is metaphysically impossible for such a state to obtain without an alternative explanation of our failure to imagine such a state. For instance, noticing that we don't have the appropriate physical makeup or noticing that we not are not acquainted with the contents of the experience in the first place, as is the case with echolocation and the like.

Haven't I thus committed myself to contradicting one of the major starting points of this debate then? Namely, haven't I implicitly rejected the idea that experience does *not* depend on the world; what Pautz called *existence neutrality?* After all, I just said that imagination is restricted by perception, and presumably perception can only be of worldly entities. I think I have not. The reason is that we can of course use our imaginative capacity to *construct* representations of entities that do not exist such as ghosts, goblins, and ghouls. Presumably, each of those examples, and any other, one could imagine, simply result from some sort of compositional process where each *bit* or *atom* does indeed occur in the world. Each property represented in those cases exists, such as color and form, but the particular combination does not. So, one can remain committed to existence neutrality, without thereby committing oneself to the claim that the constitutive features of existence-neutral experiences do not or could not co-occur.

What should one say about violations of the laws of appearance? Aren't they just compositions, in the case of the Exclusion Law, or decompositions, in the case of Berkeley's Law? I just claimed that we can remain committed to existence neutrality by utilizing some sort of composition process, and a violation to the exclusion law

seems to be just an ordinary instance of composition. I think the answer is that they *are* just instances of composition, albeit unordinary ones. That partly grounds my commitment to the claim that the Laws are not metaphysically necessary. But I think the natural thing to say here is that we are *puzzled* by the suggestion that the laws might be violated since we often have individual experiences of the constitutive components (pure-blue at location L and pure-green at location L), and we have an intuitive grasp of composition, but we can't (where can't is understood in the nomological sense) actually complete the composition. Regarding Berkeley's Law, we often experience full scenes, and we have an intuitive grasp of decomposition, but we can't (again, in the nomological sense) imagine the entire scene being decomposed (e.g., we can't imagine color without form). From a common-sense perspective, it is reasonable to think that since one can experience each of the percepts (or the whole scene), and one knows that experiences can be combined generally (or decomposed), then our failure to imagine the combination (or decomposition) might express the limits of possible experience. From the perspective of common-sense, then, it seems reasonable to assume that such an experience is impossible since if it were possible, I should be able to imagine it. The consequence of such a line of reasoning is the intuition that experiences like those are impossible. Now, my own view is that this failure is not explained by metaphysically deep facts, but by nomological facts about the visual system.

3.4. Interim Summary

In this section I have argued that perception and imagination share two things in common: (1) the same format, and (2) some of the same underlying neural

mechanisms. Specifically, in virtue of these shared properties, we can see why we can't imagine violations of the laws of appearance. It is my contention that this limitation on imagination best explains why one has the intuition that the laws of appearance are metaphysically necessary. Rather than thinking we have insight into the real nature of subjective experience, a hypothesis for which there is little evidence or independent reason to believe, we simply have a capacity to imagine possible experiences that is itself constrained, but wholly contingent. If what I say is true, then we can see why representationalism is not threatened. The reason is that there are physical constraints on imagination and perception that are absent in the case of thought. However, those physical constraints are not themselves metaphysically necessary. So, it is at least conceivable that there are perceptual states of other creatures or humans in unusual circumstances that are not so constrained, even if they are unimaginable.

4. Necessitism After All?

In the following section I want to briefly outline another possible defense of Necessitism about the laws that is in the literature. Although one might agree with me that our intuitions are not a good guide to facts about the real nature of experience, one might instead suggest that surely our intuitions might be a good guide to truths about experience in a highly restricted sense. Perhaps one thinks there is some plausible defense of the metaphysical necessity of the laws that does not depend on our intuitions about experience in general, but rather on our intuitions about our own experience. Surely, I know something about what *I* could experience. And supposing most of us are fairly similar, perhaps I also have a grasp of what *you* could experience

too. How should one proceed? On this view, one might naturally restrict the predicate "conscious visual experience" to cases of "normal human conscious visual experience". The idea underwriting this move is that if the structure or format that explains experience were different, we wouldn't be talking about *normal visual experience* anymore, but something else, call it *abnormal visual experience*. There is an analogy in the vicinity: Water is necessarily H20, and if we found something else that resembles water, but isn't H20, then it isn't water, it's #water. (See Putnam, 1962) This is the tactic taken in a recent piece by Mark Sainsbury (2022). Indeed, he claims that "…it is necessary that an experience would not be a visual experience of normal human vision unless it satisfied the laws of appearance" (pp.5-6.) The claim is that for all cases of "normal human vision", the laws of appearance hold with metaphysical necessity. Sainsbury recasts the laws as such:

> **Exclusion law****: Necessarily, a normal human visual experience does not represent a surface as pure blue and pure green at the same time.
>
> **Berkeley's law****: Necessarily, a normal human visual experience does not represent something as having a color quality without also representing it as taking up space. (p.3)

The recast laws entail that if we found a creature or a state that violated the laws, it would not be a normal human or normal visual state.

Now, I'm very skeptical of this sort of move since the predicate "normal human vision" sounds very gerrymandered. For instance, one might think that if "abnormal human vision" shares many of the same properties as experience, but without being constrained by the laws of appearance (in virtue of either having a

different physical basis or format), then it seems ad hoc to call it anything but visual experience. Of course it would be different to a certain extent, but this is not a problem in principle. After all, we are willing to allow some differences in the case of animal minds in so far as they are roughly similar. Determining the natural property that is picked out by our concept "visual experience" is beyond the scope of this paper, but presumably it is one that does not pick out humans rather than say, chameleons or Martians. Part of the justification for this is the seeming continuity between human minds and animal minds. When one is interested in studying visual experience *as such,* it would be strange to limit one's investigation to ordinary humans. So, even if the format of human vision, its architectural properties, or its neural mechanisms are sharply distinguished from other actual or merely hypothetical creatures, it doesn't follow that *visual experience* is sharply divided.

For the sake of argument, I will accept that "normal human vision" picks out a particular and interesting property since many might hesitate to call experiential states which violate the laws genuine visual states. What ought we to say about that possibility? Well, one thing to note is that it is entirely unclear that it poses a threat to representationalism either. The reason is that the best account of "normal human vision" is that it is explained by particular physical mechanisms or a particular format. And moreover, as Sainsbury rightly points out, when you restrict the predicate to normal human vision you are simultaneously rigidly designating a particular representational format. Relative to *that* format, the laws *might* follow (thought see footnote 11 above). But then we have a neurocomputational (and ultimately physical) explanation of the constraints on experience, namely, those

responsible for normal human vision. Crucially, those explanations are fully compatible with representationalism understood as a physicalist thesis of mind and experience. So, representationalism is not threatened either.

In this sub-section I presented Sainsbury's (2022) suggestion that the metaphysical status of the laws might be defended by restricting our target explanandum to normal human vision. The idea is that to have "normal human vision", one *necessarily* (in the metaphysical sense), needs to be constrained by the laws (Sainsbury, 2022.) I then argued that even if this were appropriate, representationalism need not be threatened since the best explanation of the restrictions on experience are derived from the features and properties of one's representations. This is fully compatible with there being other creatures for whom the laws do not apply. It just means that those creatures are not part of the extension of the relevant predicate (normal human vision). So, even if Necessitism about the laws of appearance regarding normal humans is true, it does not threaten representationalism about experience more generally.

This sort of defense of the modal status of the laws (as metaphysically necessary) also seems to provide an answer to one of the Necessitist's concerns. Recall that part of the problem with Necessitism about the laws of appearance is that we ought to expect a general explanation of the laws. If we restrict experience to normal human vision then we get a general explanation of the laws, i.e., they are all explained by facts about the normal human visual system.

5. Conclusion

To sum up: in this paper I have defended reductive representational about consciousness against a new and important objection known as "the laws of appearance". I first argued that there is no good reason to think the laws of appearance are metaphysically necessary. Or at least, that we have no more reason to believe in the metaphysical necessity of the laws rather than the mere nomological necessity of the laws. Importantly, if my argument works, representationalism (at least of the physicalist variety) is not undermined. I then presented Sainsbury's (2022) claim, that the laws of appearance, when restricted to normal human vision, are plausibly metaphysically necessary. However, I argued that if one accepts such a view, then representationalism is not threatened either since normal human visual experience is restricted by physical facts about the human visual system just as the extension of "water" is restricted by the fact of being composed of H20. Since any plausible representationalist thesis is presumably compatible with physical constraints on experience, one can accept the laws of appearance while maintaining the core representationalist thesis. The laws of appearance do put pressure on representationalism, but do not undermine it.

Bibliography

Alvarez, G.A. & Cavanagh, P. (2004) The capacity of visual short-term memory is set both by visual information load and by number of objects, *Psychological Science*, 15 (2), pp. 106–111. doi: 10.1111/j.0963-7214.2004.01502006.x

Aru, J. & Bachman, T. (2013). Phenomenal awareness can emerge without attention. *Frontiers in Behavioral Neuroscience*, 7, art. 891.

Aydede, M., & Güzeldere, G. (2005). Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness. *Noûs*, 39:197–255. https://doi.org/10.1111/j.0029-4624.2005.00500.x

Aydede, M. (2019). Is the experience of pain transparent? Introspecting phenomenal qualities. *Synthese*, *196*(2), 677–708. http://www.jstor.org/stable/45096372

Baars, Bernard J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.

Bae, G.-Y. & Luck, S.J. (2019). Reactivation of previous experiences in a working memory task, *Psychological Science*, 30 (4), pp. 587–595. doi: 10.1177/ 0956797619830398

Barwich, A. S. (2019). A Critique of Olfactory Objects. *Frontiers in psychology*, *10*, 1337. https://doi. org/10.3389/fpsyg.2019.01337

Barwich, A. S. (2020). *Smellosophy: What the Nose Tells the Mind*. Cambridge, MA: Harvard University Press

Batty, C. (2010). A Representational Account of Olfactory Experience. *Canadian Journal of Philosophy*, *40*(4), 511–538. DOI: https://doi.org/10.1080/00455091.2010.10716733

Batty, C. (2014). Olfactory Objects. In S. Biggs, D. Stokes, & M. Matthen (Eds.), *Perception and its Modalities.* (pp. 222–245). Oxford University Press

Bays, P.M. & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision, *Science*, 321 (5890), pp. 851–854. doi: 10.1126/ science.1158023

Bayne, T., & M. Montague (eds), (2011) *Cognitive Phenomenology*. Oxford, 2011; online edn, Oxford Academic, 19 Jan. 2012., https://doi.org/10.1093/acprof:oso/9780199579938.001.0001

Beck, J., (2012). The generality constraint and the structure of thought. *Mind*, 124: 563–600.

Beck, J. (2018). Analog mental representation. *Wiley Interdisciplinary Reviews: Cognitive Science*, 9(6), e1479.

Beck, Jacob. (2019). Perception is Analog: The Argument from Weber's law. *The Journal of Philosophy*. 116. 319-349. 10.5840/jphil2019116621.

Bharti, A. K., Yadav, S. K., & Jaswal, S. (2020). Feature binding of sequentially presented stimuli in visual working memory. *Frontiers in psychology*, 11, 33.

Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, 18: 227–47.

Block, N. (2008). Consciousness and cognitive access. *Proceedings of the Aristotelian Society*, 108 (1pt3):289-317.

Block, N. (2010). Attention and Mental Paint. *Philosophical Issues*, 20: 23-63. https://doi.org/10.1111/j.1533-6077.2010.00177.x

Block, N. (2011). Perceptual consciousness overflows cognitive access, *Trends in Cognitive Science*, 15, pp. 567–575.

Block, N. (2014). Rich conscious perception outside focal attention. *Trends in Cognitive Sciences*, 18, pp. 445–447.

Block, Ned & Stalnaker, Robert (1999). Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review* 108 (1):1-46.

Boghossian, P. & Velleman, J.D. (1989). Color as a secondary quality. *Mind* 98:81-103.

Boyer, J., Harrison., S. & Ro, T. (2005). Unconscious processing of orientation and color without primary visual cortex. *Proceedings of the National Academy of Sciences*, 102 (46), pp. 16875–16879. doi: 10.1073/pnas.0505332102

Brady, T.F., Störmer, V.S. & Alvarez, G.A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli, *Proceedings of the National Academy of Sciences* USA, 113 (27), pp. 7459–7464.

Breitmeyer, B.G., Ömen, H. & Chen, J. (2004). Unconscious priming by color and form: Different processes and levels, *Consciousness & Cognition*, 13, pp. 138– 157.

Breitmeyer, B.G., Ro, T. & Singhal, N.S. (2004). Unconscious color priming occurs at stimulus- not percept-dependent levels of visual processing, *Psychological Science*, 15, pp. 198–202.

Breitmeyer, B.G., Ro, T, Ömen, H. & Todd, S. (2007). Unconscious, stimulus-dependent priming and conscious, percept-dependent priming with chromatic stimuli, *Perception & Psychophysics,* 69, pp. 550–557.

Brentano, F. (1874). *Psychology From an Empirical Standpoint*. Routledge.

Brogaard, B., & Gatzia, D.E. (2017). Unconscious imagination and the mental imagery debate. *Frontiers in psychology*, 8, 799.

Bronfman, Z., Brezis, N., Jacobson, H. & Usher, M. (2014) We see more than we can report: 'Cost free' color phenomenality outside focal attention. *Psychological Science*, 25, pp. 1394–1403.

Burge, T. (2009). Perceptual Objectivity. *Philosophical Review*, *118*, 285–324. https://doi. org/10.1215/00318108-2009-001

Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press

Burge, T. (2014). Perception: Where mind begins. *Philosophy*, *89*(3), 385e403. https://doi.org/10.1017/S003181911400014X

Carruthers, P. (2015). *The centered mind: What the science of working memory shows us about the nature of human thought.* OUP Oxford.

Carruthers, P. (2017). Block's overflow argument. *Pacific Philosophical Quarterly,* 98, 65-70.

Carruthers, Peter (2019). *Human and Animal Minds: The Consciousness Questions Laid to Rest.* Oxford University Press.

Carvalho, F. (2014). Olfactory Objects. *Disputatio*, 6 (38):45–66

Cattaneo, Z., Vecchi, T., Pascual-Leone, A., and Silvanto, J. (2009). Contrasting early visual cortical activation states causally involved in visual imagery and short-term memory. *European Journal of Neuroscience*. 30, 1393–1400. doi: 10.1111/j.1460-9568.2009.06911.x

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2: 200–19.

Chalmers, D.J., (1996). *The Conscious Mind*, Oxford: Oxford University Press.

Chalmers, D.J., (2004). 'The Representational Character of Experience', in B. Leiter (ed.), *The Future for Philosophy*, Oxford: Oxford University Press.

Chemero, T. (2001). Dynamical explanation and mental representations. *Trends in Cognitive Sciences* 5 (4):141-142.

Chomsky, N. (1959). A review of BF Skinner's Verbal behavior. *Language*, 35 (1), 26–58.

Cichy, R.M., Heinzle, J., & Haynes, J.D. (2012). Imagery and perception share cortical representations of content and location. *Cerebral cortex (New York, N.Y. : 1991)*, *22*(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Cohen, J. (2015). 'Perceptual constancy'. In: M. Matthen (ed.), *The Oxford handbook of philosophy of perception*, 621–639, Palgrave Macmillan.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press

Cohen, M.A., Alvarez, G.A. & Nakayama, K. (2011). Natural scene perception requires attention. *Psychological Science*, 22 (9), pp. 1165–1172. doi: 10.1177/0956797611419168

Cohen, M.A. & Dennett, D.C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences,* 15, pp. 358–364.

Cohen, M.A., Dennett, D.C. & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences*, 20, pp. 324–335.

Corbetta, M. & Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, pp. 201–215.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, pp. 87–114.

Crane, T. (2003). 'The Intentional Structure of Consciousness', in Smith & Jokic (eds.), *Consciousness: New Philosophical Perspectives*, OUP.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384-398.

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. Viking.

Ding, S., Meng, L., Han, Y. & Xue, Y. (2017). A review on feature binding theory and its functions observed in perceptual process. *Cognitive Computation*, 9, pp. 194–206.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge MA. MIT Press.

Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. Cambridge: MIT Press

Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.

Fazekas, P. & Nemeth, G. (2018) Dream experiences and the neural correlates of perceptual consciousness and cognitive access. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* 373 (1755), 20170356. doi: 10.1098/rstb.2017.0356

Feigenson, L. & Halberda, J. (2008). Conceptual knowledge increases infants' memory. *Proceedings of the National Academy of Sciences*, 105, pp. 9926– 9930.

Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press

Fougnie, D. & Marois, R. (2009). Attentive Tracking disrupts feature binding in visual working memory. *Visual Cognition*, 17 (1–2), pp. 48–66. doi: 10.1080/ 13506280802281337

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). doi: 10.1167/11.12.3

Fournel, A., Sezille, C., Licón, C., Sinding, C., Gerber, J., Ferdenzi, C., Hummel, T., & Bensafi, M. (2017). Learning to name smells increases activity in heteromodal semantic areas: Learning to Name Smells. *Human Brain Mapping*, *38*, https://doi.org/10.1002/hbm.23801

Gennaro, R.J. (2018). *Higher-Order Theories of Consciousness*. Internet Encyclopedia of Philosophy. Retrieved on April 2nd, 2022 from: https://iep.utm.edu/higher-order-theories-of-consciousness/

Gibson J.J. (1979/2015). *The Ecological Approach to Visual Perception.* New York, NY: Psychology Press.

Gottfried, J.A., & Dolan, R.J. (2003). The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron*, *39*(2), 375–386. https://doi.org/10.1016/s0896-6273(03)00392-1

Gross S. (2018) Perceptual consciousness and cognitive access from the perspective of capacity-unlimited working memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373 (1755), 20170343. doi: 10.1098/rstb.2017.0343

Hahn, I., Scherer, P.W., & Mozell, M.M. (1994). A mass transport model of olfaction. *Journal of theoretical biology*, 167(2), 115–128. https://doi.org/10.1006/jtbi.1994.1057

Hardman, K.O. & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 41 (2), pp. 325– 347. doi: 10.1037/xlm0000031

Harris, J.A., Clifford, C.W., & Miniussi, C. (2008). The functional effect of transcranial magnetic stimulation: signal suppression or neural noise generation?. *Journal of Cognitive Neuroscience*, 20(4), 734-740.

Herz, R.S., & von Clef, J. (2001). The Influence of Verbal Labeling on the Perception of Odors: Evidence for Olfactory Illusions? *Perception*, *30*(3), 381–391. https://doi.org/10.1068/p3179

Horowitz, T.S. & Cohen, M.A. (2010). Direction information in multiple object tracking is limited by a graded resource. *Attention, Perception, & Psycho-physics*, 72 (7), pp. 1765–1775.

Jackson, F. (1982). Epiphenomenal Qualia, *Philosophical Quarterly*, 32: 127–36.

Jackson, F. (2004). "Mind and illusion." In P. Ludlow, Y. Nagasawa and D. Stoljar (eds.), *There's Something about Mary: Essays on the Knowledge Argument*. Cambridge, MA: MIT Press.

Jackson-Nielsen, M., Cohen, M.A. & Pitts, M.A. (2017). Perception of ensemble statistics requires attention, *Consciousness and Cognition*, 48, pp. 149–160. doi: 10.1016/j.concog.2016.11.007

Kaufman, E.L., Lord, M.W., Reese, T.W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498–525.

Keller, A. (2016). *Philosophy of Olfactory Perception*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-33645-9

Klatt, L., Getzmann, S., Begau, A. & Schneider, D. (2020). A dual mechanism underlying retroactive shifts of auditory spatial attention: Dissociating target- and distractor-related modulations of alpha lateralization, *Scientific Reports*, 10, 13860. doi: 10.1038/s41598-020-70004-2

Koch, C. & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes, *Trends Cognitive Science*, 11, pp. 16–22.

Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate.* The MIT Press.

Kosslyn S.M., Thompson W.L., Kim I.J., Alpert N.M. (1995) Topographical representations of mental images in primary visual cortex. *Nature*. Nov 30; 378(6556):496-8. doi: 10.1038/378496a0.

Kosslyn, S.M., Thompson, W. L., & Alpert, N.M. (1997). Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *NeuroImage*, *6*(4), 320–334. https://doi.org/10.1006/nimg.1997.0295

Kosslyn, S.M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J.P., Thompson, W.L., Ganis, G., Sukel, K.E., & Alpert, N.M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* (New York, N.Y.), 284(5411), 167–170. https://doi.org/10.1126/science.284.5411.167

Kriegel, U. (2003). Consciousness as intransitive self-consciousness: Two views and an argument. *Canadian Journal of Philosophy* 33 (1):103-132.

Lande, K.J. (2018). The Perspectival Character of Perception. *Journal of Philosophy.* 115 (4):187-214.

Lande, K. (2021). Mental structures. *Noûs*. 55. 10.1111/nous.12324.

Lamme, V. (2010). How neuroscience will change our view on consciousness, *Cognitive Neuroscience,* 1, pp. 204–220.

LaRocque, J.J., Lewis-Peacock, J.A. & Postle, B.R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention, *Frontiers in Human Neuroscience,* 8, art. 5. doi: 10.3389/fnhum.2014.00005

Lee, A.Y. (2021). Modeling Mental Qualities. *The Philosophical Review*. 130 (2):263-209.

Levine, J. (1983). Materialism and Qualia: the Explanatory Gap, *Pacific Philosophical Quarterly*, 64: 354–61.

Levine, J. (2003). 'Experience and Representation', in Smith & Jokic (eds.), *Consciousness: New Philosophical Perspectives*, OUP.

Lewis-Peacock, J.A., Drysdale, A.T., Oberauer, K. & Postle, B.R. (2012) Neural evidence for a distinction between short-term memory and the focus of attention, *Journal of Cognitive Neuroscience*, 24 (1), pp. 61–79. doi: 10.1162/ jocn_a_00140

Li, W., Moallem, I., Paller, K.A., & Gottfried, J.A. (2007). Subliminal smells can guide social preferences. *Psychological science*, *18 12*, 1044–1049

Luck, S. & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions, *Nature*, 390, pp. 279–281. doi: 10.1038/36846

Lycan, W.G. (1987). *Consciousness*. MIT Press.

Lycan, W.G., (2001). 'The Case for Phenomenal Externalism', in J.E. Tomberlin (ed.), *Metaphysics* (Philosophical Perspectives, Vol. 15), Atascadero: Ridgeview Publishing.

Lycan, W. (2019). Representational theories of consciousness. *Stanford Encyclopedia of Philosophy*, Retrieved on April 1st, 2022 from https://plato.stanford.edu/cgibin/encyclopedia/archinfo.cgi?entry=consciousness-representational

Ma, W.J., Husain, M. & Bays, P.M. (2014.) Changing concepts of working memory, *Nature Neuroscience*, 17 (3), pp. 347–356. doi: 10.1038/nn.3655

Macpherson, F. (2006). Ambiguous Figures and the Content of Experience. *Noûs*, 40: 82-117. https://doi.org/10.1111/j.0029-4624.2006.00602.x

Mack, A. & Clarke, J. (2012). Gist perception requires attention, *Visual Cognition*, 20, pp. 300–327. doi: 10.1080/13506285.2012.666578

Mack, A. & Manohar, S.G., Zokaei, N., Fallon, S.J., Vogels, T.P. & Husain, M. (2019). Neural mechanisms of attending to items in working memory, *Neuroscience & Biobehavioral Reviews*, 101, pp. 1–12. doi: 10.1016/j.neubiorev.2019.03.017.

Mack, A. & Rock, I. (1998). *Inattentional Blindness*, Cambridge, MA: MIT Press.

MacEvoy SP, Epstein RA. (2007). Position selectivity in scene- and object-responsive occipitotemporal regions. *Journal of Neurophysiology*, Oct;98(4):2089-98. doi: 10.1152/jn.00438.2007. Epub 2007 Jul 25. PMID: 17652421.

Mainland, J., & Sobel, N. (2006). The sniff is part of the olfactory percept. *Chemical senses,* 31(2), 181–196. https://doi.org/10.1093/chemse/bjj012

Masciari, C. F. (2021). Additional Resources for Sparse Theories of Phenomenal Consciousness. *Journal of Consciousness Studies*, 28(11-12), 125-147.

Masciari, C.F. (2022). Stop and smell the what? Two kinds of olfactory representation. *Synthese* 200, 338 (2022). https://doi.org/10.1007/s11229-022-03836-2

Masciari, C.F., Carruthers, P. (2021). Perceptual awareness or phenomenal consciousness? A dilemma. *Biology & Philosophy* 36, 18, https://doi.org/10.1007/s10539-021-09795-1

Masse, N.Y., Yang, G.R., Song, H.F., Wang, X.-J. & Freedman, D.J. (2019). Circuit mechanisms for the maintenance and manipulation of information in working memory, *Nature Neuroscience*, 22, pp. 1159–1167. doi: 10.1038/ s41593-019-0414-3

McGann, J. (2015). Associative learning and sensory neuroplasticity: How does it happen and what is it good for? *Learning & Memory*, *22*, 567–576. https://doi.org/10.1101/lm.039636.115

Millar, B. (2019). Smelling objects. *Synthese*. https://doi.org/10.1007/s11229-017-1657-8

Millikan, R.G. (1984). *Language, thought, and other biological categories: New foundations for realism*. Cambridge: MIT Press

Mongillo, G., Barak, O. & Tsodyks, M. (2008). Synaptic theory of working memory, *Science*, 319, pp. 1543–1546. doi: 10.1126/science.1150769

Morquecho-Campos, P., Larsson, M., Boesveldt, S., & Olofsson, J. K. (2019). Achieving Olfactory Expertise: Training for Transfer in Odor Identification, *Chemical Senses*, Volume 44, Issue 3, March Pages 197–203, https://doi.org/10.1093/chemse/bjz007

Naccache, L. (2018) Why and how access consciousness can account for phenomenal consciousness, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373 (1755), 20170357. doi: 10.1098/rstb.2017.0357.

Nagel, Thomas (1979). *Mortal Questions*. Cambridge University Press.

Nikolaev, A.R. & van Leeuwen, C. (2019). Scene buildup from latent memory representations across eye movements, *Frontiers in Psychology*, 9, art. 2701. doi: 10.3389/fpsyg.2018.02701

Noe, A. (2004). *Action in Perception*, Cambridge, MA: MIT Press.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of cognitive neuroscience*, *12*(6), 1013–1023. https://doi.org/10.1162/08989290051137549

Odegaard, B., Chang, M., Lau, H. & Cheung, S.-H. (2018). Inflation versus filling-in: Why we feel we see more than we actually do in peripheral vision, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373, 20170345. doi: 10.1098/rstb.2017.0345

Olson, I.R. & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the 'strong-object' hypothesis, *Perception & Psychophysics*, 64 (7), pp. 1055–1067. doi: 10.3758/bf03194756

Orlandi, N. (2011). Ambiguous figures and representationalism. *Phenomenology and the Cognitive Sciences* 10 (3):307-323.

Overgaard, M. (2018). Phenomenal consciousness and cognitive access, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373, 20170353. doi: 10.1098/rstb.2017.0353

Panksepp, J. (1998). *Series in affective science. Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press

Pautz, A. (2017). 'Experiences are representations: An empirical argument'. In B. Nanay (ed.), *Current Debates in Philosophy of Perception* (pp. 23–42). New York: Routledge.

Pautz, A. (2020). The puzzle of the laws of appearance. *Philosophical Issues*. 30: 257– 272. https://doi.org/10.1111/phis.12184

Peacocke, C. (1983). *Sense and Content: Experience, Thought, and their Relations*. Oxford: Oxford University Press.

Peacocke, C. (1986). Analogue content. *Proceedings of the Aristotelian Society*, 60: 1–17.

Phillips, I. (2015). No watershed for overflow: Recent work on the richness of consciousness, *Philosophical Psychology*, 29, pp. 1–14. doi: 10.1080/09515089.2015.1079604

Prinz, J. (2012). *The conscious brain.* Oxford University Press.

Putnam, H. (1962). It ain't necessarily so. *Journal of Philosophy*. 59 (22):658-671.

Quine, W. V. O. (1948). On What There Is. *Review of Metaphysics*. 2 (5):21-38.

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. *NeuroImage*, *50*(2), 818–825. https://doi.org/10.1016/j.neuroimage.2009.11.084
Rey, G. (1998). A Narrow Representationalist Account of Qualitative Experience. *Philosophical Perspectives*, *12*, 435–457.

Richardson, L. (2013). Sniffing and smelling. *Philosophical Studies*, *162*, 401–419. https://doi.org/10.1007/s11098-011-9774-6

Rose, N.S., LaRocque, J.J., Riggall, A.C., Gosseries, O., Starrett, M.J., Meyering, E.E. & Postle, B.R. (2016). Reactivation of latent working memories with trans-cranial magnetic stimulation. *Science*, 354, pp. 1136–1139. doi: 10.1126/ science.aah7011

Rosenthal, D.M. (2005). *Consciousness and Mind.* Oxford University Press UK.

Rupert, R.D. (2018). Representation and mental representation, *Philosophical Explorations*, 21:2, 204-225, DOI: 10.1080/13869795.2018.1477979

Sainsbury, M. (2022). Visual Experience and the laws of appearance. *Erkenntnis*. 1-8. 10.1007/s10670-021-00484-9.

Schneegans, S. & Bays, P.M. (2019). New perspectives on binding in visual working memory. *British Journal of Psychology*, 110, pp. 207–244. doi: 10.1111/bjop.12345

Sergent, C. (2018). The offline stream of conscious representations. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373, 20170349. doi:10.1098/rstb.2017.0349.

Shapiro, L. (2007). The embodied cognition research programme. *Philosophy compass,* 2(2), 338-346.

Shea, N. (2018). *Representation in cognitive science*. Oxford: Oxford University Press

Simons, D. J., & Chabris, C. F. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, 28(9), 1059–1074. https://doi.org/10.1068/p281059

Skrzypulec, B. (2018). Olfactory Objecthood. *Philosophia*. https://doi.org/10.1007/s11406-018-0017-3.

Slagter, H.A., Lutz, A., Greischar, L.L., Francis, A.D., Nieuwenhuis, S., Davis, J.M. & Davidson, R.J. (2007). Mental training affects distribution of limited brain resources. *PLoS Biology*, 5 (6), e138. doi: 10.1371/journal.pbio.0050138

Smith, B. C. (2017). Human Olfaction, Crossmodal Perception, and Consciousness. *Chemical Senses*, *42*(9), 793–795. https://doi.org/10.1093/chemse/bjx061

Speaks, J. (2017). Reply to critics. *Philosophy and Phenomenological Research*, 95, 492–506. https://doi.org/10.1111/phpr.12415

Spence, C., & Frings, C. (2019). Multisensory feature integration in (and out) of the focus of spatial attention. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s1341401901813-5

Sperling, G. (1960). The information available in brief visual presentations, *Psychological Monographs*, 74, pp. 1–29.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(5), 1565–1572.

Stein, T., Zwickel, J., Ritter, J., Kitzmantel, M. & Schneider, W.X. (2009). The effect of fearful faces on the attentional blink is task dependent. *Psychonomic Bulletin & Review,* 16 (1), pp. 104–109. doi: 10.3758/PBR.16.1.104

Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies,* 127 (1):109-166.

Stokes, M.G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework, *Trends in Cognitive Sciences*, 19, pp. 394–405. doi: 10.1016/j.tics.2015.05.004

Teghtsoonian, R., Teghtsoonian, M. (1982). Perceived effort in sniffing: The effects. *Perception & Psychophysics* 31, 324–329. https://doi.org/10.3758/BF03202655

Tootell, R. B., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy to recognition: fMRI in human visual cortex. *Trends in cognitive sciences*, *2*(5), 174–183. https://doi.org/10.1016/s1364-6613(98)01171-1

Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5

Treisman, A., and Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology. Human Perception and Performance*. 16, 459–478. doi: 10.1037//0096-1523.16.3.459

Treisman A. (1998). Feature binding, attention and object perception. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 353(1373), 1295–1306. https://doi.org/10.1098/rstb.1998.0284

Treisman, A., and Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition.* 34, 1704–1719. doi: 10.3758/BF03195932

Trübutschek, D., Marti, S., Ueberschär, H. & Dehaene, S. (2019). Probing the limits of activity-silent non-conscious working memory, *Proceedings of the National Academy of Sciences USA*, 116 (28), pp. 14358–14367. doi: 10.1073/pnas.1820730116

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of experimental psychology Human perception and performance*, *31 1*, 80–91

Tye, M., (1992). 'Visual Qualia and Visual Content', in T. Crane (ed.), *The Contents of Experience*, Cambridge: Cambridge University Press.

Tye, M. (1995). *Ten problems of consciousness,* Cambridge, MA: MIT Press.

Tye, M. (2000). *Consciousness, color and content.* Cambridge, MA: MIT Press.

Tye, M. (2020). Filling in and the nature of visual experience. *Harvard Review of Philosophy,* 27:59-69.

Usher, M., Bronfman, Z., Talmor, S., Jacobson, H. & Eitam, B. (2018). Consciousness without report: Insights from summary statistics and inattention 'blindness', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373. doi: 10.1098/rstb.2017.0354.

Ward, E.J., Bear, A. & Scholl, B.J. (2016). Can you perceive ensembles without perceiving individuals? The role of statistical perception in determining whether awareness overflows access. *Cognition*, 152, pp. 78–86. doi: 10.1016/j.cognition.2016.01.010

Ward E.J. (2018). Downgraded phenomenology: How conscious overflow lost its richness, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373 (1755), 20170355. doi: 10.1098/rstb.2017.0355

Weiskrantz, L. (1990). Outlooks for blindsight: Explicit methodologies for implicit processes: The Ferrier Lecture, *Proceedings of the Royal Society B*, 239, pp. 247–278.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung*.; 4:301–350. (Translated extract reprinted as "Laws of organization in perceptual forms." In W. D. Ellis (Ed.), (1938). A source book of Gestalt psychology (pp. 71–94). London, U. K.: Routledge & Kegan Paul Ltd.)

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to smell: Olfactory perception from neurobiology to behavior*. Johns Hopkins University Press

Wolff, M.J., Jochim, J., Akyürek, E.G. & Stokes, M.G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, 20, pp. 864–871. doi: 10.1038/nn.4546

Wood, J. N. (2013). Newborn chickens generate invariant object representations at the onset of visual object experience. *Proceedings of the National Academy of Sciences*, 110(34), 14000–14005. https://doi.org/10.1073/pnas.1308246110

Wood, J. N., & Wood, S. M. (2018). The Development of Invariant Object Recognition Requires Visual Experience With Temporally Smooth Objects. *Cognitive Science*, *42*, 1391–1406. https://doi. org/10.1111/cogs.12595

Wu, C.C. & Wolfe, J.M. (2018). A new multiple object awareness paradigm shows that imperfect knowledge of object location is still knowledge. *Current Biology*, 28 (21), pp. 3430–3434. doi: 10.1016/j.cub.2018.08.042